CRANFIELD UNIVERSITY

Daniel J.C. Skinner

A Novel Approach for Identifying Uncertainties within Environmental Risk Assessments

School of Applied Sciences

Doctor of Philosophy

Supervisor: Dr. Sophie A. Rocks

October 2012

[This page is intentionally left blank]

CRANFIELD UNIVERSITY

SCHOOL OF APPLIED SCIENCES
DOCTOR OF PHILOSOPHY

Academic Year 2012 - 2013

Daniel J.C. Skinner

A Novel Approach for Identifying Uncertainties within Environmental Risk Assessments

Supervisor: Dr. Sophie A. Rocks

October 2012

This thesis is submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

[This page is intentionally left blank]

Abstract

Uncertainties can manifest within the different aspects of environmental risk assessments, affecting the validity of the risk estimate and, in turn, weakening the basis for risk management actions. This research investigated the issues associated with uncertainty characterisation and identification in environmental risk assessments. This led to the creation of a defensible typology of uncertainties, and the creation and validation of a novel uncertainty identification system (UnISERA), based on the elicited views of experts regarding the levels (i.e. magnitudes), natures (i.e. reason for existence) and locations (i.e. where manifest) of uncertainties present within different risk domains.

The developed typology, drawn from an analysis of existing assessments, contained seven locations of uncertainty (data, language, system, extrapolation, variability, model and decision), with 20 related sub-types. The output from UnISERA, based on 19 aggregated elicitations across three risk domains (genetically modified higher plants, particulate matter and pesticides), showed that: the risk characterisation phase of assessments contained the highest magnitudes of uncertainty (the level dimension); uncertainties across all four phases of assessments existed primarily through a combination of lack of knowledge and randomness (the nature dimension); and data uncertainty was dominant in the first three phases, and extrapolation uncertainty in the final phase (the location dimension). In comparing the output from UnISERA to similarly produced results in the risk domain of engineered nanomaterials, the nature of uncertainty showed the highest degree of validation (90%), followed by the location (80%) and level (55%) dimensions.

The novel approach to uncertainty characterisation and identification presented here will be of use during environmental risk assessments and uncertainty analyses, promoting an understanding of potential uncertainties, and allowing risk analysts to perform assessments with prioritised uncertainties in mind.

Word count (text only): 44,054

Acknowledgements

Cranfield is unique, that's obvious to anyone who has studied or worked there. In my opinion, this uniqueness extends far beyond academic enrichment. To the friends gained along the way, thanks for the morning coffee-and-muffin-athons, the (select) words of wisdom and (plentiful) witticisms that only troubled research students would value. The process would have been arduous at best without you all.

To my supervisor, Sophie, who's enthusiasm for the pursuit of rigor and quality, even through extremely tough times, has been unwavering, a special thank you. I have learned many things under your guidance, though a love of Doctor Who is not one of them.

I would also like to acknowledge the valuable insight offered by Simon Pollard, Sue White, Fred Coulon and Gill Drew, as well as the anonymous experts who gave up their time to participate and without whom this research would not have been possible in its present form. I have been lucky enough to communicate my research to international audiences, made possible by grants from the Society for Risk Analysis and Cranfield Nano (namely Rob Dorey), who's support I very much appreciate.

Whilst academic development should be seen as a valuable asset, the process of self-teaching can only benefit an individual so much. There are two people who taught me 99% of what I know and value, who used all they have and all that they are to ensure that myself, my brother and my sisters could get the most from life, day after day after day. No parents could be more supportive or do more to ensure the happiness of their children. Thank you.

This process has been very enjoyable and rewarding, but has, at times, been demanding and all-consuming. During testing times we all need a source of motivation, someone from whom energy can be drawn. Most importantly, we need perspective; to be reminded of the wonderful things in our lives. To Aimée, thank you for being that endless source. The very least I can do is dedicate this work to you.

List of contents

[This page is intentionally left blank]

List of tables

List of figures

List of abbreviations

Al	Aleatory uncertainty
ALARP	As low as reasonably practicable
AM	Adaptive management
ANOVA	Analysis of variance
BBN	Bayesian belief network
BPJ	Best professional judgement
CI	Confidence interval
CDF	Cumulative distribution function
Co	Combined uncertainty
Dat	Data uncertainty
De	Deterministic uncertainty
Defra	Department for Environment, Food and Rural Affairs
Dec	Decision uncertainty
DNA	Deoxyribonucleic acid
EC	Effective concentration
EcoRA	Ecological risk assessment
EE	Expert elicitation
ENM	Engineered nanomaterial
Ep	Epistemic uncertainty
EP	Error propagation
ERA	Environmental risk assessment
Ext	Extrapolation uncertainty
FDC	Further data collection
FL	Fuzzy logic
GIS	Graphic information system
GMHP	Genetically modified higher plant
Ig	Recognised ignorance uncertainty
HHRA	Human-health risk assessment
IQR	Inter-quartile range
Lan	Language uncertainty
LC	Lethal concentration
LHS	Latin hypercube sampling
LOAEL	Lowest observable adverse effect level
LOE	Line of evidence
MCDA	Multi-criteria decision analysis
MCS	Monte-Carlo simulation
MF	Membership function
Mod	Model uncertainty
NOAEL	No observable adverse effect level
NUSAP	Numeral, unit, spread, assessment, and pedigree
PBA	Probability bounds analysis

PEC	Predicted environmental concentration
PM	Particulate matter
PNEC	Predicted no-effects concentration
PP	Precautionary principle
REACH	Registration, evaluation, authorisation and restriction of chemicals
SA	Sustainability assessment
Sc	Scenario uncertainty
ScA	Scenario analysis
SCENIHR	Scientific Committee on Emerging and Newly Identified Risks
SeA	Sensitivity analysis
SEA	Strategic environmental assessment
SI	Stakeholder involvement
SOE	Strength of evidence
St	Statistical uncertainty
Sys	System uncertainty
S-P-R	Source-pathway-receptor
TSCA	Toxic Substances Control Act
UnISERA	Uncertainty identification system for environmental risk assessments
UF	Uncertainty factor
UMT	Uncertainty management technique
UNCED	United Nations Conference on Environment and Development
US EPA	United States Environmental Protection Agency
Var	Variability uncertainty
WOE	Weight of evidence

List of publications, presentations and awards

Publications

- Skinner DJC, Rocks SA and Pollard SJT. A review of uncertainty in environmental risk: locations, levels, and options for management. Journal of Risk Research, *accepted.*
- Skinner DJC, Rocks SA, Drew GH and Pollard SJT. Identifying uncertainty in environmental risk assessments: a novel typology and implications for risk characterisation. Human and Ecological Risk Assessment, *accepted.*
- Skinner DJC, Rocks SA and Pollard SJT. A novel approach for identifying uncertainty in environmental risk assessments. *In preparation*.

Presentations

- Skinner DJC. Identifying uncertainties within environmental risk assessments. *Oral* presentation at: The Society for Risk Analysis 3rd World Congress on Risk, Sydney, 17-20 July 2012.
- Skinner DJC. A transferrable system for identifying uncertainties within environmental risk assessments. *Oral* presentation at: 31st Annual Meeting of the Society for Risk Analysis, Charleston, 4-7 December 2011.
- Skinner DJC. Developing a transferrable system for identifying uncertainties within environmental risk assessments. *Oral* presentation at: Environment Doctoral Training Centre Research Student Conference, Cranfield University, 23 November 2011.
- Skinner DJC. An evidence-based categorisation of uncertainty for environmental risk assessments. *Oral* presentation at: Risk Centre Risk and Evidence Conference, Cranfield University, 22 June 2011.

Awards

- *Travel Award* (US$2,500) from SRA secretariat for World Congress on Risk 2012, Sydney, 17-20 July 2012.
- *Student Merit Award for Best Continuing Research* (Decision Analysis and Risk Specialty Group) at the 31st Annual Meeting of the Society for Risk Analysis, Charleston, 4-7 December 2011.
- *Travel Award* (US$500) from SRA secretariat for the 31st Annual Meeting of the Society for Risk Analysis, Charleston, 4-7 December 2011.
- *Best Oral Presentation* at the Environment Doctoral Training Centre Research Student Conference, Cranfield University, 23 November 2011.

[This page is intentionally left blank]

Chapter 1: Thesis introduction

1.1 Introduction

Environmental risk assessments (ERAs) are a common feature of the environmental regulatory and decision-making process, helping to inform both short- and long-term strategy and policy development (Pollard 2001). Uncertainties are inherent to ERAs and can lower confidence in the risk estimate, in turn weakening the basis for risk management actions. Risk analysts recognise that ERAs should explicitly consider uncertainty (Funtowicz and Ravetz 1990; Costanza *et al.* 1992; Handmer *et al.* 2001). However, ERAs often fail to identify uncertainties (EEA 2007; Hart *et al.* 2007; Dale *et al.* 2008). If uncertainties are not identified, they cannot be managed (Refsgaard *et al.* 2007). If uncertainties remain unmanaged, risk estimates, and the end users' confidence in them, may be unjustified.

1.2 Aims and objectives

The aim of this research is to identify the issues associated with uncertainty characterisation and identification within ERAs, and to provide a novel approach for addressing the identified issues. The main objectives of this research are to:

i. evaluate critically the primary methods for characterising and identifying uncertainty (i.e. typologies) in environmental risk-based systems, also considering their application to ERAs;

ii. create an evidence-based typology that draws from the existing set of peer-reviewed ERAs and addresses the issues raised in objective (i), whilst also investigating the connections between uncertainty and other aspects of the ERA process;

iii. elicit experts' views on the types and magnitudes of uncertainty present within empirical (i.e. evidence-heavy) risk domains, applying the developed typology from objective (ii), leading to the creation of a generic uncertainty identification system that is organised by the different stages and tasks within an ERA; and

iv. validate the generic uncertainty identification system against the elicited views of experts in an emerging (i.e. evidence-light) risk domain, highlighting areas of strength and weakness.

This research is required in order to better acknowledge the issues associated with uncertainty identification in ERAs. The research will further help to ensure that existing uncertainty management techniques (quantification, reduction, removal; Janssen *et al.* 2003; van der Sluijs *et al.* 2004; Refsgaard *et al.* 2007) are applied reliably and commensurately. It is intended that the output from this research will be of use in the formative stages of quantitative ERAs and uncertainty analyses, and that it will promote an understanding of the potential uncertainties, helping practitioners design and perform assessments with them firmly in mind.

1.3 Thesis organisation

The thesis begins, in Chapter 2, with a literature review of uncertainty in environmental risk domains. Chapter 3 analyses existing uncertainty characterisations and introduces a summary typology – consisting of seven main locations of uncertainty across five levels – specifically for use with ERAs. Chapter 4 builds on and extends Chapter 3 by presenting a novel categorisation of uncertainties based, for the first time, on the appraisal of a large evidence base of ERAs, and further explores the relationships between identified uncertainties and other aspects of the ERA process. Chapter 5 implements the findings from the previous chapters in the form of an expert-driven approach for identifying uncertainties in ERAs. Results from three distinct risk domains are analysed and aggregated to form a single generic uncertainty identification system for environmental risk assessments (UnISERA), which is then validated against an emerging risk domain in Chapter 6. Finally, Chapter 7 contains the thesis summary and conclusions, including a discussion about the significance and limitations of the work, and options for future research.

Chapter 2: Literature review of uncertainty in environmental risk

2.1 Introduction

Government and regulatory decision-making is at the heart of protecting and managing environmental concerns. The correct implementation and enforcement of appropriate policy, through a variety of mechanisms, can ensure the long-term sustainability of natural environments. Moreover, the threats posed to humans and ecological assets (in an environmental context) can be minimised through preventative governance and efficient decision-making. These decisions, to a greater or lesser extent, all involve risk, which can be quantified and used as a basis for selecting strategies (Defra 2011). Furthermore, there is a need and a requirement to acknowledge uncertainties in order to justify the decision and build confidence.

This chapter outlines the important principles of environmental risk, introduces the tools used to quantify risk levels (risk assessments), and explores the different aspects of uncertainty in the context of environmental risk-based domains and the techniques used to manage them.

2.2 Hazard and risk

2.2.1 Defining hazard and risk

Hazard and risk are often used interchangeably (Fairman *et al.* 1998). A hazard is a chemical, physical, microbiological or psychosocial situation or property which may lead to the occurrence of harm (DHA 2002), while risk may be defined as "the combination of the probability, or frequency of occurrence, of a defined hazard and the magnitude of the consequences of the occurrence" (Royal Society 1992), or more simply as risk = likelihood x consequence (Defra 2011). A hazard can therefore be thought of as the potential to cause harm or adverse effects, and risk as the potential consequence(s) of a hazard combined with their likelihoods.

2.2.2 Risk management

Hazard-based approaches help to identify actions that can reduce sources of harm to 'safe' levels, without investigating the likelihood of that harm occurring. Targeting total or maximal 'safety' can require vast resource expenditure, often when 'safe' is either unattainable or unnecessary, because of low associated occurrence likelihoods, for example. Due to this, there has been a shift within political, economic, social, technological, and environmental regulation from using hazard-based management practices to ones that focus on risk (Fairman *et al.* 1998).

Risk management is defined as the "process of appraising options for responding to risk and deciding which to implement" (Defra 2011), and aims to "provide complete information to risk managers, specifically policymakers and regulators, so that the best possible decisions are made" (Paustenbach 1989). Risk management combines risk levels (produced during risk assessments; see Section 2.2.3) with the amount of risk that an organisation is willing to be exposed to (i.e. the risk appetite), to formulate strategies for managing the risk (Defra 2011). A good understanding of risk levels and risk appetite can provide a more solid basis for risk management actions aimed at moving risks from a position of intolerance towards one of acceptability, in the context of the UK Health and Safety Executive’s As low as reasonably practicable (ALARP) framework, for example (HSE 2001).

At its introduction in the 1970s, and for some time after, risk management processes followed a sequential approach, and included the development, evaluation, and implementation of options in response to risk (NRC/NAS 1983; Figure 2.1a).

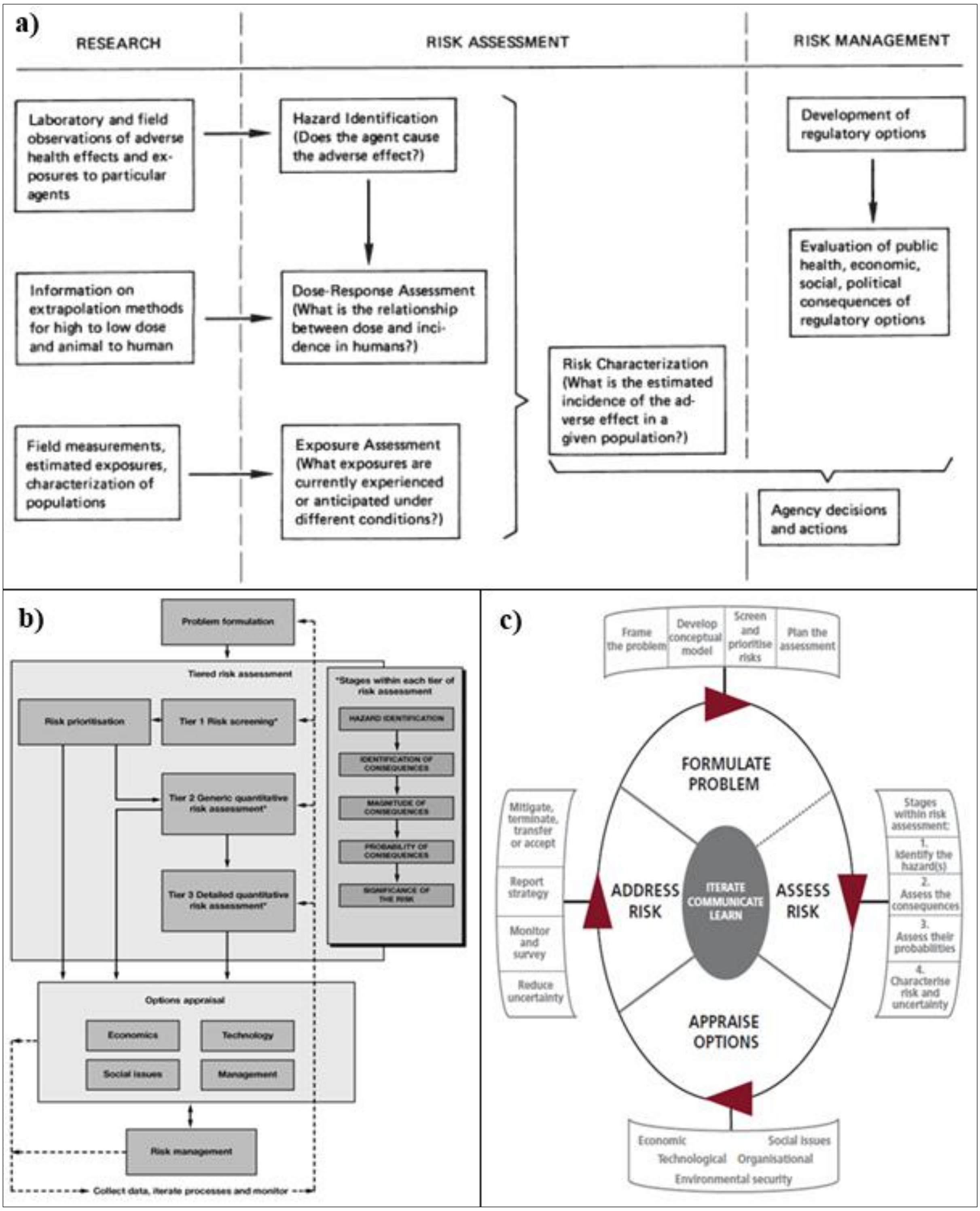


Figure 2.1 Environmental risk assessment and management through a) a sequential approach (NRC 1983), b) a tiered approach (DETR/EA and IEH 2000), and c) a cyclical approach (Defra 2011).

A tiered management approach followed, which, through its iterative form, ensured that the level of effort put into assessing each risk was proportionate to its priority and complexity (DETR/EA 2000; Figure 2.1b). More recently, a cyclical approach has been used to represent environmental risk management as a dynamic process rather than a single, one-off exercise (Defra 2011; Figure 2.1c). Throughout, the main goal of risk management has been to move risk from a position of intolerance to one of acceptability (HSE 2001), by either terminating, mitigating, transferring, exploiting, or accepting the risk (Defra 2011). For this to happen, the level(s) of risk must first be assessed.

2.2.3 Assessing risk

Numerous methods exist for assessing and communicating the levels of risk associated with different hazards, including human-health risk assessment (HHRA), strategic environmental assessment (SEA), sustainability assessment (SA), ecological risk assessment (EcoRA), and environmental risk assessment (ERA; Zhang *et al.* 2010). Within each assessment type, the level of risk can be described either qualitatively (i.e. using broad categorisations such as 'low', 'medium', and 'high', or through narrative descriptions), quantitatively (i.e. using numerical values), or semi-quantitatively (i.e. a combination of qualitative and quantitative) depending on the available information and resources and the level of output-precision required (Fairman *et al.* 1998). Selection of the most appropriate framework depends largely on the asset(s) in need of protecting, which in this research are environmental.

2.3 Environmental risk

2.3.1 Environmental risk assessments

ERAs must identify the potential sources of harm in the environment, those who may be exposed to the harm (receptors; e.g. ecosystems, animals, plants, people), and the connections between the two (pathways; DETR/EA and IEH 2000). The objective of an ERA is to investigate this source-pathway-receptor (S-P-R) paradigm in order to quantify the levels of risk posed to the receptors by the different sources of harm. In this context, the potential for harm can include naturally occurring phenomena, such as flooding and earthquakes, as well as those hazards which stem directly from human action, such as the development and application of synthetic chemicals. ERAs typically focus on one of three receptors, namely environmental assets, ecosystems, or human health, and can therefore encompass some or all of the processes typically found within HHRAs and EcoRAs. ERAs include the same four stages:

- hazard identification and problem formulation, where potential hazards are postulated and explored through one or more S-P-R paradigms;
- exposure assessment, which determines the probability that the receptor(s) will be exposed to the hazard in a given period of time;
- effects assessment, which establishes the quantity or level of the hazard required for adverse effects to ensue; and
- risk characterisation, where a level of risk is produced and evaluated in terms of its significance (US EPA 1992; Fairman *et al.* 1998; US EPA 1998; DETR/EA and IEH 2000; DHA 2002; Defra 2011).

ERAs generally use a single type of evidence in the evaluation of risk levels. However, multiple types of evidence can also be integrated into a formal assessment process, known as weight of evidence (WOE).

2.3.2 Weight of evidence assessments

WOE definitions

A critical element of the environmental decision-making process (and often the environmental risk assessment process) is the amalgamation of different types of evidence and the evaluation of the degree to which they support a conclusion (Linkov *et al.* 2009). The frameworks used in this process are termed weight of evidence and aim to provide either a definitive course of action or information as to what additional research needs to be conducted for definitive conclusions to be reached (Chapman 2007). Conclusions are formulated using different types of data (lines of evidence; LOE) that vary in the degree to which they support or refute a particular hypothesis (strength of evidence; SOE). For clarity, a LOE refers to the data and/or information that can be used to arrive at a conclusion regarding a stated hypothesis (e.g. biological, toxicological, or financial; Linkov *et al.* 2009).

Since its introduction, there has been much confusion concerning the proper meaning and use of the weight of evidence term (Doull *et al.* 1996; see Table 2.1). A common aspect of all WOE frameworks is the integration of different lines of evidence during the risk assessment process (Linkov *et al*. 2009). This integration of divergent evidence distinguishes WOE assessments from other forms of risk assessment. Non-WOE assessments may yield a risk characterisation based on a single line of evidence (e.g. biological), whilst WOE assessments produce conclusions based on at least two lines of evidence (e.g. biological and chemical).

Table 2.1 Key weight of evidence definitions from the literature.

Source	WOE definition
Menzie *et al.* (1995)	Process by which multiple measurement endpoints are related to an assessment endpoint to evaluate whether significant risk of harm is posed to the environment.
Burton *et al.* (2002)	Process of combining information from multiple lines of evidence to reach a conclusion about an environmental system or stressor.
Krimsky (2005)	Process or method in which all scientific evidence that is relevant to the status of a causal hypothesis is taken into account.
Linkov *et al.* (2009)	Framework for synthesizing individual lines of evidence, using methods that are either qualitative ... or quantitative … to develop conclusions regarding questions concerned with the degree of impairment or risk.

WOE classifications and implementation methods

WOE frameworks can be broadly categorised according to their function:

- Qualitative frameworks often do not attempt to fully integrate different lines of evidence and instead consist of systematic reviews of work relating to a topic (Linkov *et al.* 2009);
- Quantitative frameworks use mathematical models in their process and produce a numerical output (Weed, 2005); or

- Semi-quantitative frameworks that attempt to combine the positive aspects of both, and usually have quantitative methodologies and qualitative interpretations (Chapman, 2007).

Within each of these three categories, different methods exist which aim to associate a level of risk with one or more sources of potential harm. Different WOE methods are described by Chapman *et al.* (2002), Burton *et al.* (2002), Weed (2005), and Linkov *et al.* (2009; see Table 2.2).

Table 2.2 Examples of weight of evidence methods available when quantifying risk.

WOE framework type	WOE methods
Qualitative	Best professional judgement, qualitative combination, narrative review, listing evidence.
Semi-quantitative	Ranking, indexing, logic systems, causal criteria, decision matrix, broad-scale WOE, quality criteria, meta-analysis.
Quantitative	Statistics, mathematical and computational modelling, multi-criteria decision analysis.

Qualitative WOE methods

Qualitative methods, whilst simple and cost-effective to execute, are prone to bias and subjectivity and are very rarely transparent or transferrable. Best professional judgement (BPJ) is one such method (Chapman *et al.* 2002; Linkov *et al.* 2009). Also referred to as expert judgement or expert opinion soliciting, it uses the opinion of one or more experts in a field with regard to available data. For example, Efroymson and Suter (2001) use a BPJ approach in which WOE is tabulated using positive (+) and negative (-) symbols to represent the degree of risk associated with each element investigated. Whilst the methodology employed suggests that conceptual models can easily be drafted using this highly qualitative approach, it also shows that distinctions between investigated parameters can be made with little or no supporting evidence. Whilst these studies can be performed on a limited budget and in a short period of time, the assigned weightings are hugely subjective often leading to bias which can be both hard to detect and eliminate (Wang *et al.* 2007).

The qualitative combinations technique (Burton *et al.* 2002) refers to the integration of lines of evidence in a non-quantitative manner (Efroymson and Suter, 2001). The main drawback of this form of review is that it will be likely to yield different results if conducted by another body (Kavlock and Cummings 2005).

Narrative reviews (Weed 2005) consist of a review of existing literature and comment on the current state of the science or make research recommendations. A crucial element is the selection of literature; the databases and search terms used will determine the quality and relevance of the review.

The final method, listing evidence (Linkov *et al.* 2009), is not considered a true WOE framework, since no integration of evidence is performed (Linkov *et al.* 2009). For example, Oller and Erexson (2007) use existing literature to evaluate the ability of nickel sulphate hexahydrate to induce micronuclei in rat bone marrow, and attempt to associate this material to human respiratory cancer. However, the work is very suggestive since no human-based testing is used to support claims made.

Semi-quantitative WOE methods

Semi-quantitative techniques are often used for retrospective hazard assessment, but offer a consistent and systematic approach in which risk characterisation is the priority (Chapman *et al.* 2002; Linkov *et al.* 2009). Many of the identified techniques within the semi-quantitative bracket bear strong similarities to one another.

Ranking (also known as *scoring*; Chapman *et al.* 2002, Burton *et al.* 2002, Linkov *et al.* 2009), which assigns weights to lines of evidence, is usually based on BPJ. Expert ranking involves the use of BPJ (e.g. Calabrese *et al.* 1997), consensus ranking includes the opinions and weightings of stakeholders in the process (e.g. Menzie *et al.* 1996), and semi-quantitative ranking allows data to be normalised to percentiles and evaluated in tandem (e.g. Cherry *et al.* 2001). Whilst ranking techniques may be applicable to a number of conditions and environments, weightings are usually applied through qualitative means, meaning that interpretations will vary according to their spatial and temporal implementation (Burton *et al.* 2002).

Frameworks involving indices (Chapman *et al.* 2002; Linkov *et al.* 2009) assign weights to lines of evidence, and integrate the separate lines into a single index-based result. For example, Schmidt *et al.* (2002) combine physical, chemical, toxicological, and ecological parameters into a single number in the range 0 to 100. Whilst the ranking and indexing methods merge expert judgement with communicative quantification, neither technique provides a basis for the explanation of parameter weighting assignments, causing transparency and reproducibility issues (Chapman 2007).

Logic based systems (Chapman *et al.* 2002; Linkov *et al.* 2009) attempt to combat the flaws of other semi-quantitative methods by adopting standardised methods or guidelines to integrate lines of evidence. Suter *et al.* (2002) have used such a system to assess impairments in aquatic ecosystems employing BPJ, a standardised WOE analysis, and reconsideration of alternatives where no clear conclusions are formed. Categorisation of results against existing logical criteria (such as government guidelines) encourages transparency, accountability, and ease of interpretation (Weeks and Comber, 2005). However, in situations where categorisations are borderline, BPJ is often used as a selection mechanism, rather than continued quantitative research (Basketter *et al.* 2006).

Decision matrices (Burton *et al.* 2002) consist of binary alternatives (e.g. toxic or not toxic) relating to different lines of evidence. This is essentially a form of tabular ranking (e.g. Menzie *et al.* 1996), but is more robust than the previously-mentioned methods.

Broad-scale WOE frameworks (Burton *et al.* 2002) incorporate several WOE methods into one process. This generalised approach may contain qualitative, semi-quantitative, and quantitative elements (Suter *et al.* 2002), causing the underlying methods to be complex because of the number of disparate techniques involved.

Quality criteria methods (Weed 2005) are based on work by Klimisch *et al.* (1997) who address the quality of toxicological data by organising it into one of four groups: reliable without restriction; reliable with restriction; not reliable; or not assignable. Evidence considered reliable can be used in the risk assessment, while evidence considered not reliable or not assignable is not automatically included, but may be used subject to expert judgement. This approach is more likely to be influenced by subjective bias, and relates more to the quality of data used than to the risk assessment itself (Weed 2005).

The causal criteria group (Weed 2005; Linkov *et al.* 2009), based on the research of Hill (1965), describes nine criteria (consistency of association, strength of association, dose-response, temporality, experimentation, specificity, biologic plausibility, coherence, and analogy) that are used to establish evidence of causation. Conclusions drawn using this approach are specific to the researcher(s), since no standardised method is detailed describing the relative importance (weighting) of each criterion. Results are usually expressed as probabilities, making them easily understandable and comparable. However, the assignment of weightings to parameters is performed using qualitative methods, meaning that one of the most important stages of the assessment is non-transferrable (Swaen and van Amelsvoort 2009).

An extension of the causal criteria category is meta-analysis (Weed 2005) where results from several studies are combined using a common measure of the strength of relationship between two variables (e.g. Bailer *et al.* 2002). This method is less subjective than those belonging to the causal criteria category, although the relevance of the stressor-effect relationship remains a matter of judgement.

Quantitative WOE methods

Quantitative WOE methods use formalised mathematical approaches to apply weightings to parameters and thereby weigh the body of evidence (Linkov *et al.* 2009). Quantitative methods are time-intensive and require more data than qualitative and semi-quantitative methods to function effectively, though the high levels of transparency involved mean that decisions made are more defensible than through less structured approaches (Linkov *et al.* 2009).

Statistical techniques, such as statistical summarisation (Chapman *et al.* 2002) and quantitative likelihood functions (Burton *et al.* 2002) are based on the statistical testing of quantitative data. Bailer *et al.* (2002) propose a WOE system that uses p-values (which are the probabilities that an observed difference between two groups is due to chance alone) as the mechanism to quantify the level of risk. Statistical approaches are robust, transparent, and applicable to a wide range of environments, but can also be very complex and inappropriately applied, leading to false confidence in results (Burton *et al.* 2002).

Statistical techniques are commonplace; however other methods also are used. Mathematical models (Linkov *et al.* 2009) take advantage of increasing processing speed and memory capacity to implement techniques in a computational setting. Results can also be presented through the use of graphic information systems (GIS), aiding interpretation through the use of intuitive displays. However, creating models from empirical data carries the assumption that future events will occur at the same frequency and magnitude as previous events. Application of these types of models also requires a significant amount of data in order to develop probabilities, which can be an expensive and time-intensive process. However, there are many advantages of combining mathematical methods with computational power, including: transparency; efficiency; reliability, speed, and through the use of globally available data (such as topographic, demographic, and meteorological) their high level of transferability (Lee and Choi, 2004; Neuhäuser and Terhorst, 2007).

Multi-criteria decision analysis (MCDA; Linkov *et al.* 2009) is another type of quantitative tool. MCDA methods evaluate and choose among alternatives based on multiple criteria using systematic analysis (Belton and Steward 2002). Processes often involve screening, sorting, ranking, and selecting alternatives from a number of options. The pairwise techniques employed in some MCDA tools require parameter weightings, as well as upper and lower value thresholds, to be assigned normally by BPJ. Alteration of these sensitive (and possibly subjective) values can yield varying results (Yatsalo *et al.* 2007), and important distinctions between expert opinions and stakeholder values can easily become blurred along the way (Stahl *et al.* 2002). Conversely, the iterative nature of MCDA allows parameter weightings to be adjusted in order to explore their significance, leaving the body of scientific evidence unaffected. Iteration can also be used to address underperforming alternatives or drive future research and development streams (Kiker *et al.* 2008). Similarly, MCDA tools can evaluate the sensitivity of conclusions to changes in input parameters, and also allows lines of evidence to be weighed independently from social, political, and economic considerations, which may otherwise unfairly impact on results (Linkov *et al.* 2009).

ERAs that evaluate risk according to single or multiple lines of evidence are found in numerous environmental disciplines, and are often an enforced component of regulated risk domains.

2.3.3 Example environmental risk domains

National, supranational and global regulatory agencies are required to consider 'landscapes' of environmentally-focused risks. The specific risks that they must regulate vary from organisation to organisation, but largely comprise of a combination of man-made and naturally-occurring sources. The ERA structure (see Section 2.3.1) is designed to be used in association with anthropogenic sources of potential harm only (EPA 1998). Some examples of risk domains, which will also feature as case studies later on in this research (see Chapters 5 and 6), follow.

Genetically modified higher plants

The genetic modification of plants involves the manipulation of their deoxyribonucleic acid (DNA) and the inter- and intra-transference of genes across species and generations, in order to achieve the stable expression of desirable traits (Altieri 2000). The traits expressed by genetically modified higher plants (GMHPs), also termed transgenic plants, include insect- and disease-resistance, tolerance to certain pesticides (including herbicides, insecticides, and fungicides), and increased yield and nutritional quality (Wolfenbarger and Phifer 2000). Despite their present-day cultivation in numerous countries around the world, GMHPs remain a divisive issue, largely because their obvious benefits must be reconciled against the potential risks associated with their adoption. These potential risks may include (EFSA 2010):

- Persistence and invasiveness of the GMHP, including plant-to-plant gene transfer;
- GMHP-to-micro-organism gene transfer;
- Interaction of the GMHP with target organisms;
- Interaction of the GMHP with non-target organisms;
- Impact of the specific cultivation, management, and harvesting techniques;
- Effects on biogeochemical processes; and
- Effects on human and animal health.

The potential risks of GMHPs to humans and the environment have been discussed at length by academics, industrial scientists, and regulators (Auer 2008), the majority of whom promote, or enforce, the use of ERA on a case-by-case basis.

Particulate matter

Particulate matter (PM) is a mix of airborne solid particles of varying size, shape, solubility, composition, and origin (Pope III and Dockery 2006). Assessing the potential impacts that PM has on human health has been a concern of epidemiological and toxicological science for several decades (Deck *et al.* 2001). Such impacts are noted to include cardiovascular and cardiopulmonary morbidity, respiratory-based disease, the development of cancerous tumours, and even increased rates of mortality (Greene and Morris 2006). The existence and onset of these deleterious effects is heavily influenced by the frequency and duration with which the human receptor is exposed to the PM, as well as the physical size of the particles. A mass of air-suspended PM is typically categorised according to a 50% cut-point of particles at a certain aerodynamic diameter: coarse PM_{10}, with a diameter of between 10 µm and 2.5 µm, predominantly comprises dust and soil particles; fine $PM_{2.5}$, with a diameter of between 2.5 µm and 1.0 µm, is primarily the result of industrial combustion processes; and ultrafine $PM_{1.0}$, with a diameter of <1.0 µm, is also associated with combustion sources, although the particles have a short life and quickly aggregate (Pope III and Dockery 2006). Size categorisations are important in a human-health risk context, since smaller particles may be more likely to penetrate from the lungs to the bloodstream and translocate to other parts of the body (WHO 2006).

Pesticides

A pesticide is a chemical that is designed to repel or mitigate the effects of pests, such as insects, weeds, and microorganisms, through exertion of toxic action (Chèvre *et al.* 2006). Pesticides, which include insecticides, herbicides, and fungicides, are mainly used in the agricultural sector, with benefits ranging from improved yield, quality, nutritional value, and cosmetic appearance (Damalas and Eleftherohorinos 2011). However, the persistence and transport of pesticides can cause wide-ranging negative effects to a multitude of non-target receptors across different environmental compartments (Chèvre *et al.* 2006). Regulatory authorities, such as the United States Environmental Protection Agency (US EPA) and the Chemicals Regulation Directorate in the UK, recognise the potential risks of the chemicals in pesticides, and often oversee strict licensing processes in which risk assessments (environmental, ecological, and human) are key aspects (Shwarzman and Wilson 2009).

Engineered nanomaterials

Engineered nanomaterials (ENMs) are a novel class of substances that have been engineered at the molecular-level to achieve unique mechanical, optical, electrical, or magnetic properties (Chio *et al.* 2012). There has been considerable debate about the nomenclature associated with ENMs, even what constitutes an ENM (Klaine *et al.* 2008); the definition is now widely accepted as being a material with one or more dimensions in the nanoscale (< 100 nm; BSI 2007). The main advantage of ENMs is in their reduced size, which allows them to behave differently to their parent material resulting in unique properties (Renn and Roco 2006). The range of these properties, such as enhanced antibacterial, thermoelectric, and immunological performance, provides enormous benefits to a number of distinct areas of research and production (Chen *et al.* 2011). However, the rapidly increasing utilisation of ENMs has raised concerns that their release into the environment, at all stages of their life-cycle, can lead to a variety of potential risks (Quik *et al.* 2011). These risks require assessment and management (Handy *et al.* 2008).

In contrast to the domains of genetically modified higher plants, particulate matter and pesticides, the legislation concerning ENMs is still very much in development. In the EU, existing guidelines for chemical risk assessment and management, the REACH regulation (registration, evaluation, authorisation and restriction of chemicals) currently applies to ENMs (EC 2008). However, the European Commission's independent Scientific Committee on Emerging and Newly Identified Risks (SCENIHR) has indicated that the methods advised in REACH, principally associated with exposure and effects assessment, require significant development to cope with the unique challenges presented by ENMs (SCENIHR 2009).

The presence of uncertainties within ERAs, whatever the risk domain, can affect the validity of the formulated risk levels as well as the credibility of following decisions (Dale *et al.* 2008). Therefore, it is recognised by scientists and non-scientists alike that such assessments should explicitly consider uncertainty (Funtowicz and Ravetz 1990; Costanza *et al.* 1992; Handmer *et al.* 2001).

2.4 Uncertainty

2.4.1 A basic introduction to uncertainty

Uncertainty is defined in a number of distinct ways in the literature and in practice, allowing a number of distinct interpretations (Refsgaard *et al* 2007; Troldborg 2010). Protection and regulatory agencies often consider high-level definitions, for example that uncertainty "refers to our inability to know for sure" (US EPA 2010), that it concerns "known impacts and unknown probabilities" (EEA 2007), or that it reflects "limitations in knowledge about environmental impacts and the factors that influence them" (Defra 2011). In these definitions and their accompanying descriptions, uncertainty is likened to a poor state of knowledge borne through a lack of relevant information. Whilst uncertainty may exist in situations where there is missing or minimal amounts of data (Walker *et al.* 2003), it may also persist where information is complete and freely available (van Asselt and Rotmans 2002). Therefore, ascribing uncertainty to gaps in knowledge alone, as the basic definitions offered by strategically-positioned regulatory agencies seemingly do, does not provide a full description of potential uncertainties.

Research conducted in the natural and physical sciences often expresses 'uncertainty' through confidence intervals and absolute or relative error (Krayer von Krauss 2005). This implies that uncertainty is a statistical problem that can be described adequately through statistical means. However, only a small range of potential uncertainties fall within this statistical set (Walker *et al.* 2003).

Uncertainty exists in numerous forms throughout risk assessments and the wider decision-making landscape (Ascough II *et al.* 2008). Even those assessments that are clearly defined and meticulously executed need to be communicated in a level of detail and to an appropriate audience. These distinct facets represent potential sources of uncertainty. Uncertainty is far too complex a concept to be distilled into one or two sentences. However, the plethora of lengthy characterisations that do exist may promote confusion rather than clarity.

2.4.2 Uncertainty analysis

Risk analysts deal with uncertainty by using uncertainty analysis (or uncertainty assessment), which identifies, analyses and manages uncertainties within the different phases of risk assessments (ECHA 2008). Uncertainty analysis is typically performed during the risk characterisation phase, and can range from simple point descriptions of uncertainty to highly detailed probabilistic analyses. The type of analysis conducted largely depends on the complexity of the assessment in which it features. For example, a 'scoping' level assessment requires no more than a single point description of associated uncertainties, whilst a quantitative assessment demands a deterministic treatment of uncertainty at the very least (EFSA 2006). In the context of ERAs, a reliable and robust uncertainty analysis, whatever its complexity, necessitates an understanding of all potential uncertainties.

2.4.3 Environmental uncertainty

Much of the early research surrounding uncertainty in environmental domains focussed on characterising the physical flaws in acquiring experimental data (Veseley and Rasmuson 1984; Henrion and Fischoff 1986; Alcamo and Bartnicki 1987; Beck 1987). This mechanistic view was later expanded to include the way in which natural processes change over space and time, and also our ability to communicate in a clear and consistent manner (Finkel 1990; Morgan and Henrion 1990). A strategic view of uncertainty followed, and with it came the notion of significance (Wynne 1992; Faucheux and Froger 1995; Stirling 1998). More recently, these concepts have been combined, with uncertainty investigated through its constituent dimensions (Walker *et al.* 2003; Figure 2.2); namely its *location* (where the uncertainty is manifest in the system of interest), *nature* (uncertainty due to the incompleteness of knowledge or the inherent variability of natural systems), and *level* (describing the severity of the uncertainty, ranging from determinism to ignorance).

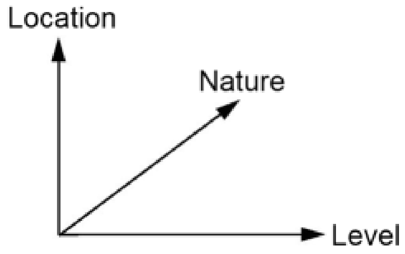


Figure 2.2 The three dimensions of environmental uncertainty: the location, where the uncertainty is manifest in the system of interest; the nature, uncertainty due to the incompleteness of knowledge or the inherent variability of natural systems; and the level, describing the severity of the uncertainty, ranging from determinism to ignorance (after Walker *et al.* 2003).

Identifying the different uncertainties that exist in applied situations is an essential part of the uncertainty management process (Morgan *et al.* 1990). The location of the uncertainty must be known, since without knowledge of its physical representation no further action can be implemented. The nature of the uncertainty dictates the degree to which it can be managed; knowledge-based uncertainties can be quantified, reduced, and potentially removed, whilst those uncertainties that reflect the randomness of natural processes can only be quantified. Finally, the severity of the uncertainty informs selection of the most appropriate management technique (Refsgaard *et al.* 2007). In order to effectively manage uncertainty, it is essential that all dimensions are first considered and identified (Walker *et al.* 2003; Janssen *et al.* 2003; Refsgaard *et al.* 2007; Knol *et al.* 2009).

2.5 The location of uncertainty

2.5.1 System understanding

In the formative stages of risk-based assessments (i.e. problem formulation, hazard identification) one or more conceptual models of the system under investigation are built to depict the various sources of harm, the likely receptors, and the ways in which these two groups are thought to be connected. Boundaries are set determining the inclusion or exclusion of important features, potentially affecting the completeness of the representation (Walker *et al.* 2003; Janssen *et al.* 2003). The potential issues associated with such processes are primarily related to a lack of understanding about the system of interest (Rowe 1994). Dewulf *et al.* (2005) argue that, in some cases, there can also be too much information, leading to multiple frames of reference to understand a phenomenon. From either position (too much or too little) the associated uncertainties, termed contextual (Walker *et al.* 2003; Janssen *et al.* 2003), conceptual (Rowe 1994), ambiguity (Dewulf *et al.* 2005), and process (Ascough II *et al.* 2008), all reflect the limits of scientific understanding.

Generally, where understanding is low, uncertainty will be high and vice-versa. System uncertainty can therefore impact ERAs wherever understanding is lacking. However, a field which develops rapidly, such as nanotechnology, may contain high levels of knowledge as well as some system-related uncertainties, due largely to the unknowns that progress brings. For example, the contribution of physical structure to a nanoparticle's toxicity may only be partly understood, whilst its effects upon different receptors of interest may simply be unknown (Zalk *et al.* 2009). Such system uncertainties can heavily impact the exposure and effects phases of ERAs.

2.5.2 Data

Whether empirical or experimental, all data carries a level of inherent confidence associated with its truth and correctness. Identifying potential sources of uncertainty can help to distinguish between the reliable and the unreliable. According to Morgan *et al.* (1990) the most common data uncertainty concerns errors in the direct measurements of a quantity. This type of uncertainty, termed random (Henrion and Fischoff 1986), statistical (Morgan *et al.* 1990; Finkel 1990), or measurement (Regan *et al.* 2002), refers to the variation across multiple measurements of the same quantity. It stems from failings in the sampling process (Finkel 1990), including operator error and instrument error (Regan *et al.* 2002). All measurable empirical quantities, such as the speed of light, will be inexact to some (*apparently* random) degree (Morgan *et al.* 1990). Whilst the magnitude of this uncertainty can easily be calculated through statistical testing of the unexplained variation in measurements or by simply obtaining more data (Henrion and Fischoff 1986), a related, but distinct, uncertainty can be harder to quantify.

Systematic uncertainty (Henrion and Fischoff 1986; Finkel 1990; Morgan *et al.* 1990; Regan *et al.* 2002) is defined as the difference between the true value of the quantity of interest and the value to which the mean of the measurements converge as the sample size increases

(Morgan *et al.* 1990; Regan *et al.* 2002). Systematic uncertainties can arise from undetected errors, in, for example, the experimental procedure (Henrion and Fischoff 1986), and from unintentional errors, such as the erroneous calibration of measuring equipment (Regan *et al.* 2002). This type of uncertainty is dealt with by detecting errors and bias in the experimental procedure and data employed and attempting to remove them. Whilst simple in theory, the interaction of multiple techniques and processes makes it much more difficult in reality (Regan *et al.* 2002). Furthermore, practitioners must exercise their own judgement, potentially allowing for high levels of user-subjectivity to dominate (Henrion and Fischoff 1986).

Separate data concerns may stem from the way in which the data are analysed and interpreted (Regan *et al.* 2002; Maier *et al.* 2008), and from incomplete or unavailable data records (Maier *et al.* 2008). For example, the same (potentially uncertain) data values may be used in a number of distinct assessments, simply because no other data are available (McColl *et al.* 2000).

In the context of ERAs, data uncertainties are most common in the analysis phase, where original experimental data are primarily used. For example, McColl *et al.* (2000) discuss the effect that a limited or erroneous data record can have on the establishment of dose-response levels for use in a contaminated site assessment, which may then be adopted in other separate assessments, potentially beyond the discipline of contaminated land. The data uncertainties directly impact on estimates of risk and, by extension, the quality of environmental decision-making (Faucheux and Froger 1995).

2.5.3 Model

Modelling is an attempt to understand processes within a system of interest, predict responses, evaluate management alternatives, and support the policy and decision-making process (Arhonditsis *et al.* 2007). The associated procedures vary, though they routinely involve an initial conceptualisation stage (see Section 2.5.1), which is then developed into a numerical and/or computational representation (Stephens *et al.* 1993). Modelling relies heavily, in one form or another, on data.

The data uncertainties associated with modelling mostly refer to the quality of the data used to populate system variables (i.e. the input data; Vesely and Rasmuson 1984; Rowe 1994), but they also extend to the model's parameters. These values, which may be exact (e.g. π), fixed (e.g. the gravitational constant g), measured *a priori*, or derived through calibration (Walker *et al.* 2003; Krayer von Kraus 2005), are the unvarying constants within a model; they are coefficients which are stable in time and space. Whether derived through direct measurement or using an empirical database for calibration, parameter uncertainty (Alcamo and Bartnicki 1987; Beck 1987; Morgan *et al.* 1990; Bedford and Cooke 2001; Huijbregts *et al.* 2001; Janssen *et al.* 2003; Walker *et al.* 2003; Maier *et al.* 2008; Ascough II *et al.* 2008) is primarily a reflection of the uncertainties associated with data (see Section 2.5.2). Furthermore, it is likely that this data will be representative of a different location, scale, and time span to the model's input variables and parameters (Troldborg 2010), forcing interpolation and/or extrapolation (similar to the data availability issues discussed previously). Parameter uncertainties are also associated with the structure of the model.

Computational and/or numerical models are simplified versions of real-world phenomena (Ascough II *et al.* 2008). The challenge for modellers is to balance the highest achievable level of realism against financial, computational, and temporal restraints. In addition to these constraints, the uncertainties associated with model representativeness, termed model structure uncertainty (Alcamo and Bartnicki 1987; Beck 1987; Janssen *et al.* 2003; Walker *et al.* 2003), model uncertainty (Finkel 1990; Bedford and Cooke 2001, Huijbregts *et al.* 2001; Regan *et al.* 2002) or method uncertainty (Maier *et al.* 2008), stem from a lack of understanding about the system. Structural uncertainties may include: the definitions of, and the physical relationships between and among, the variables and parameters (Ascough II *et al.* 2008; Knol *et al.* 2009); different views on the correct interpretation of observations and theories and their subsequent implementation (Regan *et al.* 2002); approximations in numerical solution, including rounding and precision of numerical values (van der Sluijs 1997); and the initial conceptual plans and boundary conditions adopted (Alcamo and Bartnicki 1987; see Section 2.5.1). This final concern outlines the potential importance of the relationship between the conceptual model and its physical implementation: an oversimplification of the conceptual model may result in a failure to capture essential features, leading in turn to inadequate numerical or computational simulations, whilst an undersimplification may yield a model that is too complex, and therefore financially, computationally, and temporally expensive to build and execute (El-Ghonemy *et al.* 2005). In effect, model structure uncertainties convey reservations about knowledge of the current state of a natural system, the future evolution of the system, or both (Walker *et al.* 2003).

The implementation of knowledge, in the context of the technical aspects of computational modelling, has further limitations. These uncertainties specifically relate to the software and hardware used (Rowe 1994; van der Sluijs 1997; Janssen *et al.* 2003; Walker *et al.* 2003; Ascough II *et al.* 2008). Software errors arise from bugs in developer and operational platforms, poorly-designed algorithms, and mistakes in code, to name a few (Walker *et al.* 2003). Hardware errors arise, quite simply, from bugs in the hardware (van der Sluijs 1997).

The input, parameter, structural, and technical hindrances discussed all manifest in physical models, limiting operational capability and ultimately affecting output (Walker *et al.* 2003, Janssen *et al.* 2003; Ascough II *et al.* 2008). Even those models which are good representations of the real-world, providing consistently accurate results, can never be completely exact. Following this logic, Morgan *et al.* (1990) come to the stark conclusion that "every model is definitely false", it is simply a matter of by how much.

Within ERAs, model uncertainties should be considered wherever numerical or computational models are utilised, which is principally during the analysis phase. For example, ApSimon *et al.* (2002) describe the uncertainty associated with modelling complex atmospheric processes, such as deposition, within the exposure phase of a trans-boundary air pollution ERA. Furthermore, the output from the modelling process, which may be used in part to help formulate risk estimates, should, if not already managed, be treated with due caution at the risk characterisation stage.

2.5.4 Human

Human error is a more recently acknowledged source of uncertainty (Maier *et al.* 2008). These uncertainties refer to the unintentional human-based failings in assessments that are not covered by system-knowledge, models, and data, and that are generally of a more qualitative, reflective, and interpretive character (Janssen *et al.* 2003). This may, for example, include conflicts between individuals and/or groups (disagreement uncertainty; Morgan *et al.* 1990), varying perspectives and values resulting in irreconcilable differences (value diversity; Rowe 1994; van Asselt and Rotmans 2002), or the societal importance of an individual, elevating their views above those of others (stakeholder uncertainty; Maier *et al.* 2008).

Human uncertainties can exist at any stage of ERAs, from unintentionally subjective actions at the problem formulation phase to stakeholder disagreements concerning tolerability thresholds during risk characterisation. For example, in a multi-criteria approach for prioritising sites in sediment management Alvarez-Guerra *et al.* (2009) account for the human uncertainty involved in assigning weightings to the different criteria, the result of unintentionally biased opinions brought about by past experiences. Human-based uncertainties are also strongly linked to the way in which language is used to communicate.

2.5.5 Language

The uncertainties associated with language arise for a number of reasons, but stem primarily from a lack of clarity (Morgan *et al.* 1990). Language can be used to express ideas and commands or to communicate results. Language can be controlled. Therefore, theoretically at least, the associated uncertainties can be eliminated. However, words and their connected meanings can evolve and change, making isolation and treatment more difficult in applied situations.

The components of language can be considered "partial truths" since they can be interpreted differently by different people (Li *et al.* 2006). Linguistic variables may be: ambiguous (Bedford and Cooke 2001; Regan *et al.* 2002; Ascough II *et al.* 2008), where more than one meaning can be drawn, and it is not clear which is intended; underspecific (Regan *et al.* 2002; Ascough II *et al.* 2008), where terms do not provide the level of precision required; or vague (Regan *et al.* 2002; Ascough II *et al.* 2008), where borderline cases are permitted to exist, resulting in the blurring of distinctions between terms. The use of a single field-specific term can contain all of these failings (Acosta *et al.* 2010). In addition to these, two further linguistic uncertainties have been postulated, context dependence (Regan *et al.* 2002), where there is a failure to properly convey the context in which a term is to be understood, but is rarely an issue with detailed communication, and indeterminacy of linguistic terms (Regan *et al.* 2002), which represents the unknown future developments of languages and the resulting effects on incorporated terms.

Language, in the context of ERAs, is not phase-specific. As such, the associated uncertainties should be expected to exist in various locations throughout the process, from basic definitions to the communication of risk levels (Keiter *et al.* 2009). For example, language uncertainties can easily exist within the expert elicitation exercises that are often used for information gathering, evidence-checking, or results validation (Acosta *et al.* 2010).

2.5.6 Variability

Whereas system-knowledge, models, data, human, and language uncertainties largely correlate to knowledge-based failings, variability uncertainty is concerned with the random states of systems of interest. The antithesis of the accidentally subjective human uncertainties discussed previously (see Section 2.5.4) are the practices, termed human variability uncertainty (Rowe 1994; van Asselt and Rotmans 2002), which occur from intentionally biased and subjective human actions (Khan *et al.* 2002). Humans invariably display bias when they have something to gain, and subjectivity when they believe their own views to be more correct than those of others (Chen *et al.* 2007). Human variability can be exhibited by those with close links to a project as well as those with a lower vested interest Croke *et al.* 2007).

The naturally variable aspects may be considered unexpected, but free from intentional bias (Jørgensen *et al.* 2009). The associated uncertainty, termed natural variability uncertainty (Finkel 1990; Rowe 1994; van Asselt and Rotmans 2002; Huijbregts *et al.* 2001; Regan *et al.* 2002; Ascough II *et al.* 2008) pertains to the chaotic traits of nature, i.e. the unpredictable quality of natural processes (Ascough II *et al.* 2008; Regan *et al.* 2002). Further distinction can be made between those sources of natural variability which occur across spatial scales (e.g. over a reference grid) and those which occur over temporal scales (e.g. from one year to the next; Rowe 1994; Huijbregts *et al.* 2001; Regan *et al.* 2002). Since natural variability is intrinsic to nature, it is also intrinsic to the corresponding aspects within ERAs, from factors affecting the fate and transport of a stressor in exposure assessment (Schwartz *et al.* 2000), to the difference in responses shown by receptors of the same species during effects assessment (Borsuk *et al.* 2006), to the variability in determining appropriate tolerance thresholds in risk characterisation (Chen and Ma 2007).

Two additional categories of variability are proposed. Technological variability (Van Asselt and Rotmans 2002; Ascough II *et al.* 2008) refers to the unexpected surprises that technological developments and breakthroughs bring, which are triggered by human action, but can be considered to have a stochastic quality. Institutional variability (Funtowicz and Ravetz 1990; van Asselt and Rotmans 2002; Ascough II *et al.* 2008) can be considered to be human variability that is exhibited over large groups and organisations (e.g. stakeholders and societies), and includes aspects such as social values, economic principles, and cultural dynamics (van Asselt and Rotmans 2002). With respect to ERAs, technological variability can occur wherever such systems are in place, and institutional variability is most likely to exist in assessments at the community or population scale (e.g. epidemiological studies).

2.5.7 Decision

Decision uncertainty (Finkel 1990; Ascough II *et al.* 2008), also termed volitional (Bedford and Cooke 2001) or choice-laden uncertainty (Huijbregts *et al.* 2001), exists when doubt surrounds an optimal course of action, often in the face of differing objectives (Finkel 1990). There may be a situation where multiple options satisfy at least a part of the criteria for a decision, or where no such alternative exists. These uncertainties exist within the ERA process, principally at the risk characterisation phase, but also in a wider risk management context. For example, the management of ecological and environmental resources requires decision-makers to evaluate multiple and often conflicting strategies, whilst balancing objectives of productivity and sustainability (Ducey and Larson 1999). Such decisions will be heavily influenced by the results of environmental projects, and can therefore comprise any or all of the other outlined uncertainties (Ascough II *et al.* 2008).

2.5.8 The location of uncertainty in the context of ERAs

The location in which uncertainty manifests depends on the different aspects of the system being explored. For example, an assessment of a novel potential risk (e.g. an exotic animal disease) in an open natural environment, involving multiple data, models, and stakeholders, can potentially contain all of the uncertainties discussed. When certain aspects do not feature, such as modelling processes, the related uncertainties will not be an issue. However, even the most basic assessments will be likely to include system-knowledge, data, human involvement, use of language, and variable qualities. Whilst uncertainties can manifest individually several are likely to exist, meaning that the full range of location-based uncertainties described here should be considered (Refsgaard *et al.* 2007).

2.6 The nature of uncertainty

2.6.1 Aleatory uncertainty

Aleatory uncertainty, an aspect of the nature dimension (Figure 2.2), represents the inherent randomness displayed in human and natural systems (Bedford and Cook 2001; Ascough II *et al.* 2008). Also termed physical (Vesely and Rasmuson 1984), stochastic (Helton 1994), variability (Hoffman and Hammonds 1994; Janssen *et al.* 2003; Walker *et al.* 2003; Hayes 2006), random (Bevington and Robinson 2002; Regan *et al.* 2002), or ontic (Petersen 2006; Knol *et al.* 2009), aleatory uncertainty cannot be reduced, although additional research may help to better understand the complexities of the system(s) of interest. Whilst such systems may in actuality be chaotic rather than random (and are therefore in principle understandable; Regan *et al.* 2002), risk analysts find it useful to treat the associated uncertainties from the latter position. For example, stochastic numerical techniques (such as Monte-Carlo simulation and Latin Hypercube sampling) act as realistic representations of real-world processes, which are either viewed as being too complex for deterministic interpretation (e.g. seismic activity) or as inherently random (e.g. weather systems). However, in mimicking nature, stochastic models can produce results that are consistently more representative than their deterministic counterparts (Hromkovic 2005).

2.6.2 Epistemic uncertainty

Epistemic uncertainty (Bedford and Cooke 2001; Walker *et al.* 2003; Petersen 2006; Ascough II *et al.* 2008; Knol *et al.* 2009) represents the imperfection of knowledge concerning a system of interest. Also termed completeness (Vesely and Rasmuson 1984; Rowe 1994), subjective (Helton 1994), knowledge-based (Hoffman and Hammonds 1994; Janssen *et al.* 2003), or systematic (Bevington and Robinson 2002), and in contrast to aleatory uncertainty, epistemic uncertainty can be quantified, reduced, and possibly eliminated, depending on the specific situation. However, there is an important caveat to this point. Whilst epistemic uncertainty is in principle reducible by increasing relevant knowledge, this new information can reveal the true depths of our ignorance, only serving to increase the associated uncertainty (Janssen *et al.* 2003; van der Keur 2008).

2.6.3 The nature of uncertainty in the context of ERAs

It can be difficult to distinguish between epistemic uncertainty and aleatory variability in applied ERAs. In such instances the dividing line can be blurred by problem-specific features such as the current level of subject knowledge (Janssen *et al.* 2003). This is important as it is increasingly recognised that uncertainty and variability need to be treated separately (Li *et al.* 2008; Kumar *et al.* 2009; Qin and Huang 2009), due to the differing degrees to which they can be managed (see Section 2.4.4).

2.7 The level of uncertainty

2.7.1 Understanding the level of uncertainty

Humans exhibit a variety of distinct levels of knowledge, ranging from determinism (perfect knowledge) to indeterminacy (lack of knowledge; Wynne 1992). The further one moves from a deterministic understanding of a system, the more severe the uncertainty becomes (Walker *et al.* 2003). The level of uncertainty (Figure 2.2) is specifically described according to two factors, namely the degree of confidence attached to the likelihood of an event occurring, and the degree of confidence attached to the severity of outcomes should that event occur (Stirling 1999; Figure 2.3).

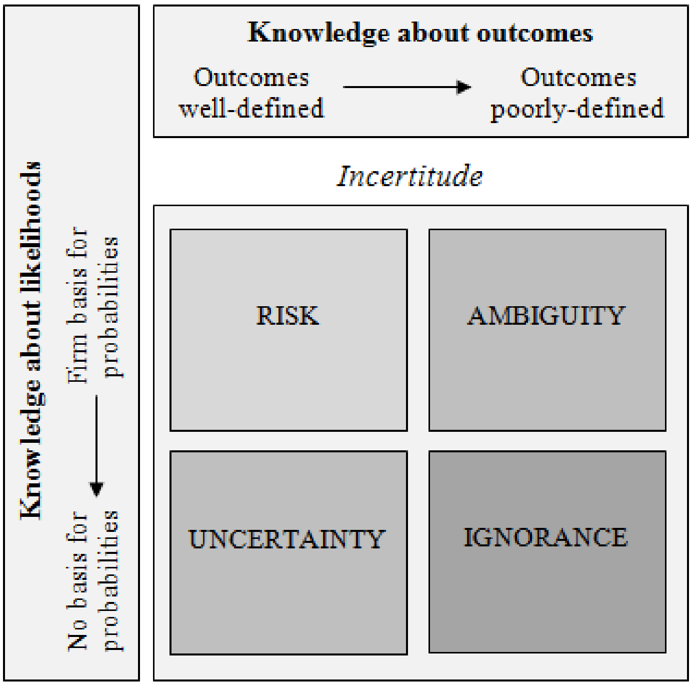


Figure 2.3 Uncertainty levels as defined through knowledge about likelihoods and knowledge about outcomes (after Stirling 1999).

The level of uncertainty can also be represented using a linear spectrum (Figure 2.4).

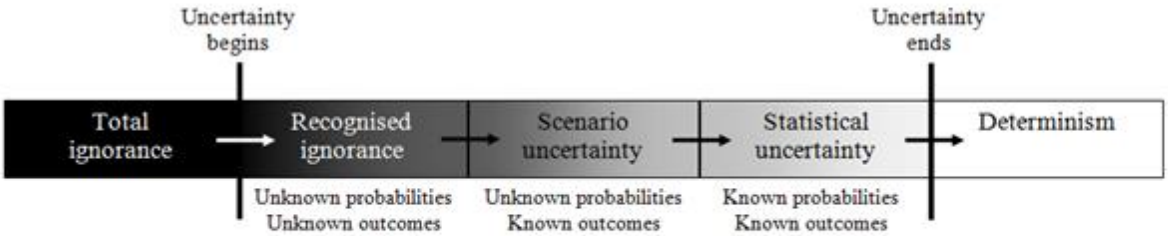


Figure 2.4 Schematic showing the spectrum of uncertainty levels, defined through knowledge about likelihoods and knowledge about outcomes (after Walker *et al.* 2003 and Krayer von Kraus 2005).

These metrics are used to convey the level of understanding, and therefore the severity of the associated uncertainty.

2.7.2 State 1: knowing a lot

At the deterministic end of the spectrum sits the range (here termed State 1) in which risk analysts are most comfortable, where uncertainty is low. First described by Funtowicz and Ravetz (1990; and later by van der Sluijs 1997 and van Asselt and Rotmans 2002), ‘inexactness’ refers to the specified events for which we “roughly know” the likelihoods and outcomes, and where significant digits and error bars are the assessment tools of choice (Funtowicz and Ravetz 1990). Due to the applicability of risk assessment methodologies in treating this level of uncertainty (e.g. using frequency distributions), the term ‘risk’ has also been applied (Wynne 1992; Stirling 1999). In this context, ‘risk’ describes a situation for which one can define a full range of possible outcomes and present their associated probabilities (Stirling 1999; Figure 2.3). Adopting this same definition, other proposed terms include ‘probabilistic’ (Beer 2006), ‘statistical’ (Brouwer and Blois 2008; Knol *et al.* 2009) and ‘certainty’ (Faucheux and Froger 1995).

2.7.3 State 2: knowing the probabilities

In moving away from determinism the degree of understanding diminishes. In doing so, one first comes to the state in which one can confidently assign probabilities to events, but have little understanding of the ramifications (State 2). Termed ‘ambiguity’ (Stirling 1999), ‘conflicting evidence’ (van Asselt and Rotmans 2002), ‘statistical’ (Walker *et al.* 2003; Janssen *et al.* 2003; Petersen 2006), ‘incertitude’ (Beer 2006), or ‘qualitative’ (Brouwer and Blois 2008), this level of uncertainty refers to a situation in which “we don’t know what we know” (van Asselt and Rotmans 2002). Statistical measures can be used to constrain likelihoods (i.e. probability distributions; Janssen *et al.* 2003) with techniques such as sensitivity analysis and fuzzy logic used to better understand the outcomes (Stirling 1999). Whilst this uncertainty is the type most often considered in the natural sciences, greater attention should be paid to the other uncertainties that are more severe (Walker *et al.* 2003).

2.7.4 State 3: knowing the outcomes

Moving farther from determinism, the next level of uncertainty (State 3) describes the state in which there is confidence about the outcomes of an event, but not in assigning likelihoods to that event occurring (i.e. the reverse of State 2). Termed ‘unreliability’ (Funtowicz and Ravetz 1990; van der Sluijs 1997), ‘uncertainty’ (Wynne 1992; Stirling 1999), ‘practically immeasurable’ (van Asselt and Rotmans 1999), or ‘ambiguity’ (Beer 2006), it refers to the position in which “we know what we do not know” (van Asselt and Rotmans 1999). Since probabilities escape recognition, typical risk assessment methodologies cannot be applied, and the range of possible outcomes must instead be used to describe associated uncertainties. The term ‘scenario’ is also used when referring to this state (Walker *et al.* 2003; Janssen *et al.* 2003; Brouwer and Blois 2008; Knol *et al.* 2009), because of a reliance on the analysis of scenarios when attempting to resolve probabilities.

2.7.5 State 4: knowing a little

If there neither exists a basis to define probabilities nor a complete set of outcomes, one moves into a state of ‘ignorance’ (State 4; Wynne 1992; Faucheux and Froger 1995; Stirling 1999; van Asselt and Rotmans 2002; Brown 2004; Beer 2006), and it becomes necessary to proceed with due caution (Stirling 1999). The definitions associated with this term do not consider the difference between the ignorance of which one is aware and that of which one is totally oblivious. The terms ‘borderline ignorance’ (Funtowicz and Ravetz 1990; van der Sluijs 1997) and ‘recognised ignorance’ (Walker *et al.* 2003; Janssen *et al.* 2003; Brouwer and Blois 2008; Knol *et al.* 2009) are also used, since, by definition, “we cannot say anything useful about that of which we are ignorant”. The ideal solution is to increase knowledge of the problem, thus reducing uncertainty, and move back towards determinism (Walker *et al.* 2003).

2.7.6 State 5: not knowing

Being aware of uncertainty and having some related evidence is a serious dilemma requiring much attention. However, it is not as serious as being altogether unaware of uncertainty. The inverse of deterministic knowledge is ‘indeterminacy’ (State 5; Wynne 1992; van Asselt and Rotmans 2002; Brown 2004). This is the deepest and most important form of uncertainty, since it is the uncertainty of which one knows nothing (Walker *et al.* 2003). The origins and effects of an event can only be observed once the event has occurred, at which point a transition towards a state of awareness (i.e. further towards determinism) can take place.

2.7.7 The level of uncertainty in the context of ERAs

Determining the level of uncertainty helps to focus attention toward the features of ERAs that are most uncertain, and, when approached with the nature and location, allows selection of the most appropriate managing tool(s; Refsgaard *et al.* 2007). Resolving the level also allows the uncertainty to be described in an appropriate manner. For example, using statistical measures to describe uncertainties closest to the indeterminacy end of the spectrum is inappropriate because nothing is known of the associated statistical distributions (Krayer von Kraus 2005).

Once identified, the uncertainties within the three dimensions should be managed using one or more available techniques (Janssen *et al.* 2003; Refsgaard *et al.* 2007).

2.8 Uncertainty management techniques

2.8.1 Adaptive management

Adaptive management (AM) aims to reduce uncertainty in the decision-making process, and attempts to incorporate the needs of scientists, managers, and other stakeholders into a system where differing alternatives and objectives are present (Dey *et al.* 2000). It does this through an iterative approach, consisting of stages such as problem definition, design, implementation, monitoring, evaluation, and adjustment (Williams *et al.* 2009). In this way, AM can be thought of as a process of continual learning and improvement.

In research concerning the adverse effects of cooling water intake structures, Dey *et al.* (2000) propose the use of AM in combination with a widely accepted ecological risk assessment framework. Resource managers can reduce this uncertainty by adopting flexible AM techniques that can be modified when new stock-based information becomes available. The combination of an ecological risk assessment process with AM ensures that risks are identified and investigated, and that uncertainty is highlighted and dealt with throughout.

2.8.2 Bayesian belief networks

A Bayesian belief network (BBN) is a graphical representation of a system, in which uncertain characteristics are represented by distinct nodes, and causal associations between nodes are represented by joining lines (Jensen 1996). The relationships between all possible states of each node and all other nodes to which it is linked are expressed through probability values, which may be probabilistic, discrete, or subjective, depending on the data used (Aspinall *et al.* 2003). Once a BBN has been constructed, it can be executed quickly and efficiently to calculate probabilities for each state of interest, and is therefore an efficient decision-making tool.

2.8.3 Bootstrapping

Bootstrapping is a combined statistical and computational method where a parameter in a simulation is resampled (and replaced) using the corresponding value from an original dataset (King and Richardson 2003). A number of separate iterations are performed, each time with a different parameter value. The aim of bootstrapping is to identify relevant thresholds for particular parameters; these levels are the points at which a change in the value brings about a noticeable change in the effect. For example, in an investigation into the subsea release of oil from a riser, Nazir *et al.* (2008) use a bootstrapping technique consisting of 10,000 iterations to identify the 95th percentile of a parameter, along with its associated uncertainty.

2.8.4 Confidence intervals

A confidence interval is often used to estimate the reliability of a parameter, or the distribution range in which that parameter is likely to feature. Most commonly represented as a percentage, the interval expresses the level of confidence in the estimate made. Confidence intervals are directly associated with percentiles, such that a 95% confidence interval is analogous to the 95th percentile of the data in question (Nazir *et al.* 2008).

2.8.5 Error propagation

Error propagation describes the effect of carrying uncertainties from individual parameters through to the functions with which they are built (Thorsen *et al.* 2001). It attempts to quantify the effect that uncertainty in the input variables has in the confidence of the final result(s). The uncertainty (or error) can be described in a number of ways, but is most commonly represented as a relative error, absolute error, or standard deviation (e.g. Finnveden *et al.* 2009).

2.8.6 Expert elicitation

Expert elicitation is a structured method for describing and quantifying uncertain aspects of systems from the viewpoint of relevant experts (Slottje *et al.* 2008; US EPA 2009; Knol *et al.* 2010). It is primarily employed in situations where required information is insufficient or unavailable (Knol *et al.* 2010), with experts drawing from related empirical evidence and theoretical insight (US EPA 2009). Expert elicitation became commonplace with the advent of the Delphi method (Brown *et al.* 1969; Dalkey 1969; Rowe and Wright 1999). Since then, numerous other elicitation protocols have been described (Kahneman *et al.* 1982; Morgan *et al.* 1990; Cooke 1991; Meyer and Booker 1991), and range from face-to-face group sessions to remotely-conducted questionnaires targeted at individuals. Whichever protocol is applied, a central domain of interest, against which the views of experts can be elicited, must be established.

2.8.7 Further data collection

The collection of increased quantities of data are an ideal way to enhance knowledge and understanding as well as, hopefully, to reduce uncertainties. Situations in which data are sparse, imprecise, or uncertain in some other way, will benefit greatly from the assemblage of more information. Those processes that build upon the underlying data, such as modelling or decision-making, should in turn have their associated uncertainties reduced. For example, Avagliano and Parrella (2009) highlight the importance of identifying key parameters for which more precision and accuracy is needed in order to reduce uncertainty in model output. Confidence in the data are paramount: increasing confidence in the data, upon which everything else is based, reflects positively upon the reliability of the risk assessment as a whole.

2.8.8 Fuzzy logic

Derived from fuzzy set theory, fuzzy logic is a form of multi-valued logic that allows its components to be approximate rather than precise. Introduced by Zadeh (1965), fuzzy logic offers an alternative to binary logic in which elements either "belong" or "not belong" to a set. Fuzzy sets allow its constituents to belong to a set, but only to a certain degree as defined by its membership function (MF), which is a continuous function between [0,1] (Acosta *et al.* 2010).

Fuzzy sets may take one of two forms: type-I, where the MF of a set member is crisp; or type-II, where the MF of a fuzzy-set member is itself fuzzy, meaning that each element within a type-II system has a MF which is type-I rather than crisp (Zadeh 1975). This added fuzziness improves the handling of uncertainties relating to the MF of an element, and avoids the problem of defining an exact MF when it is not straightforward to do so (Acosta *et al.* 2010).

The fuzzy logic systems (FLS) which build on fuzzy sets comprise of a series of if-then rules (Wang 1997), and generally contain three parts: a fuzzifier, which turns crisp inputs into fuzzy inputs; an inference block, which processes the fuzzy inputs (according to the if-then inference rules) to produce fuzzy output sets; and a de-fuzzifier, which transforms the fuzzy sets into a crisp output (Acosta *et al.* 2010).

Fuzzy logic, a flexible uncertainty handling technique both in terms of implementation and of intuitive understanding (Kaloudis *et al.* 2005), is able to handle data imprecision (Acosta *et al.* 2010), and also provides a methodology for computing directly with words (Zadeh 1996).

2.8.9 In situ data collection

This simple coping mechanism advises the collection of required data directly from the study site. It attempts to negate the uncertainty associated with extrapolating data and observations from the laboratory into a field environment, or from one field location to another. Oughton *et al.* (2008) postulate that collecting site-specific data can help to reduce uncertainties associated with intrinsic local variability, but also that it can be a time- and resource-intensive process. In situ collection may be considered similar

2.8.10 Latin hypercube sampling

Latin hypercube sampling (LHS) is a stratified technique based on the subdivision of a probability density function (PDF) into distinct intervals; the number of segments created is equal to the number of samples required for use as input to a numerical model, and therefore also equal to the number of model executions performed (Klier *et al.* 2008). Each uncertain input parameter has an associated PDF from which random samples are drawn, similar to the MCS approach. However, representative segmentation of the PDF in LHS ensures that the upper and lower ends of the distributions used in the analysis are well represented (Helton and Davis 2003).

LHS is considered to be more efficient than simple random sampling, since it requires fewer simulations to produce the same precision (Helton and Davis 2003). As such, LHS is often used within MCS processes in order to reduce the number of runs required.

2.8.11 Monte-Carlo simulation

Monte-Carlo simulation (MCS) is a computational technique that utilises repeated executions of numerical models to simulate stochastic processes (Qin and Huang 2009). Application of the procedure requires that at least one of the input parameters to a model be uncertain, and that the uncertainty be represented as a probability density function (PDF; US EPA 1996). Each execution of the model employs a distinct value for the parameter in question, as selected randomly from the PDF. The model results can then be combined to construct a further PDF which indicates the risk estimate (Ma 2002). MCS effectively allows users to quantify uncertainty in model estimates as a function of input parameter uncertainty (US EPA 1996), and is the principal technique used in mainstream software such as @RISK (Palisade Corporation, Ithaca, NY) and Crystal Ball (Oracle Corporation, Redwood Shores, CA).

MCS is traditionally used to quantify levels of uncertainty attached to model inputs, but it is also capable of computing the associated variability. To characterise both the uncertainty and variability, a PDF is first created for each for the parameter under investigation (Eckhardt

1987). A nested execution is then performed, where the uncertainty is handled in the first loop and the variability in the second (Eckhardt 1987). The outcome of this process, known as two-dimensional or second-order Monte-Carlo, is a collection of cumulative distribution functions (CDFs) that display both the uncertainty and variability in the results (Wu and Tsang 2004).

2.8.12 Multi-criteria decision analysis

MCDA is a technique used when faced with multiple and often conflicting decision alternatives. MCDA methods bring together criteria and performance scores, usually in matrix form, to provide a basis for integrating risk and uncertainty levels. In this way, it is possible to perform an evaluation (ranking) of the alternatives (Linkov *et al.* 2007). The main advantage of MCDA is its capacity to draw attention to areas of conflict between stakeholders and decision-makers. These participants may initially be unready to relinquish their own subjective views, but through the use of MCDA a deeper understanding of the values held by others is possible (Critto *et al.* 2007; see Section 2.3.2).

2.8.13 No action

The recognition of uncertainty should be the catalyst for an effort to deal with it. However, the quicker, simpler, and cheaper option is to do nothing. This can be considered scientifically unsound, but may be the only way to proceed when faced with constraints of time, money, or knowledge. On other occasions though, there may be little excuse for identifying an uncertainty and not attempting to quantify, reduce, or remove it. Application of this mechanism may lead to undesirable consequences and will probably be difficult to justify to stakeholder groups (Oughton *et al.* 2008).

2.8.14 Precautionary management

This type of management is based upon the application of the Precautionary Principle (PP), as defined during the United Nations Conference on Environment and Development (UNCED; also known as the Earth Summit) in Rio de Janeiro, 1992. Principle 15 of this declaration (termed the Rio Declaration) states that:

"In order to protect the environment, the precautionary approach shall be widely applied by States according to their capabilities. Where there are threats of serious or irreversible damage, lack of full scientific certainty shall not be used as a reason for postponing cost-effective measures to prevent environmental degradation." (UN 1992).

The PP is used as a basis for taking mitigating action in the presence of uncertainty; it has been developed to guide human activities with unpredictable risks (Godduhn and Duffy 2003), and therefore provides for the occurrence of events rather than for a reduction in the uncertainty associated with them.

2.8.15 Probability density function

A PDF is a technique used to account for parametric uncertainty, and is often associated with model inputs (Ciffroy *et al.* 2009). The PDF function of a parameter describes the relative likelihood that it will occur at a specific point; it is represented as a distribution in graph form and describes the frequency of occurrence for different parameter values over a given range. Once created, a PDF may be used in the modelling process to relate the possible values of a parameter to the likelihood that they will be observed in the real system (Stephens *et al.* 1993). Each execution of the model would therefore invoke a new selection from the appropriate PDF.

PDFs not only give information regarding the most probable value of a parameter, and can thus help guide proper selection during the modelling process, but also provide a range of all potential values (Ciffroy *et al.* 2009). PDFs can also be used for statistical investigations outside computational modelling, but are predominantly used as a basis for sampling techniques such as MCS or LHS.

2.8.16 Sensitivity analysis (intra and inter)

Sensitivity analyses are fast and straightforward approaches for dealing with uncertainty and are primarily associated with assessments involving computational modelling processes (Huysmans *et al.* 2006). A sensitivity analysis examines the contribution of uncertainty associated with input parameter values to the endpoint of interest (Oughton *et al.* 2008). Techniques involved effectively test the sensitivity of a chosen output variable to variation in quantities relating to input variables (Huysmans *et al.* 2006). Several sensitivity analysis methods exist, including sample (Pearson) and rank (Spearman) correlation, analysis of variance (ANOVA), and classification and regression tree (CART; Chen *et al.* 2007). Selection of an appropriate method is defined by such factors as computing power and time available, the quantity of uncertain parameters being modelled, and the relationship between the input parameters and the endpoints of interest (Oughton *et al.* 2008).

The literature identifies two distinct forms of sensitivity analysis: intra, where an analysis is performed using a single model (e.g. Chen *et al.* 2007; Oughton *et al.* 2008); and inter, where an analysis is performed using results from two or more models (e.g. Liao and Chou 2005; Arhonditsis *et al.* 2007). Selection of the appropriate method is dependent upon model availability and specific requirements.

2.8.17 Uncertainty factors

An uncertainty factor (UF) may be thought of as a margin of safety: it attaches a factor-based correction to the data being used which reflects the level of uncertainty within it (Phillips *et al.* 2008). According to the US EPA (2002) UFs are used to address five forms of uncertainty: interspecies variability, predominantly when extrapolating from animals to humans; intraspecies variability, either between humans or other species; when extrapolating from a lowest observable adverse effect level (LOAEL) to a no observable adverse effect level (NOAEL); database uncertainty, which concerns the use of incomplete and/or unreliable records; and extrapolation from less than lifetime exposures to lifetime exposures. Recommendations are also made concerning the numeric value of the factors to be applied in different situations. For example, a factor of 10 is consistently applied where data are extrapolated from animal-specific studies to represent human-specific endpoints (Calabrese *et al.* 1997; US EPA 2002). This level asserts that humans are to be considered 10-fold more sensitive than animals to the hazard under investigation (Phillips *et al.* 2008).

2.9 Conclusion

Risk assessments are common features of the environmental decision-making process. However, associated uncertainty can affect the validity of the formulated risk levels as well as the credibility of following decisions. Environmental uncertainty is considered to comprise three dimensions, namely its *location* (where the uncertainty is manifest in the system of interest), *nature* (uncertainty due to the incompleteness of knowledge or the inherent variability of natural systems), and *level* (describing the severity of the uncertainty, ranging from determinism to ignorance), with each dimension containing a number of sub-types of uncertainty. The phases of ERAs contain a variety of distinct tasks and processes and have the potential to contain some or all of the wide range of uncertainties discussed in this chapter. ERAs should identify and manage these uncertainties using one or more existing techniques, to ensure decision confidence. For this, a reliable characterisation of potential uncertainties, both in the context of wider risk-based domains and ERAs, is required.

Chapter 3: A discursive analysis of environmental uncertainty typologies

3.1 Introduction

The identification of uncertainties within ERAs often relies on risk analysts considering lists of potential uncertainties. These lists, commonly termed uncertainty typologies, aim to define and communicate the important features of uncertainty. Uncertainty typologies are useful in helping practitioners better understand the associated concepts (Morgan and Henrion 1990) and identify uncertainties (Knol *et al.* 2009). Some risk analysts will make use of a single uncertainty typology, but it is more likely that several versions will be used (van Asselt and Rotmans 2002). Uncertainties cannot be managed until they have been identified, and they cannot be identified until the potential types are understood (Figure 3.1).

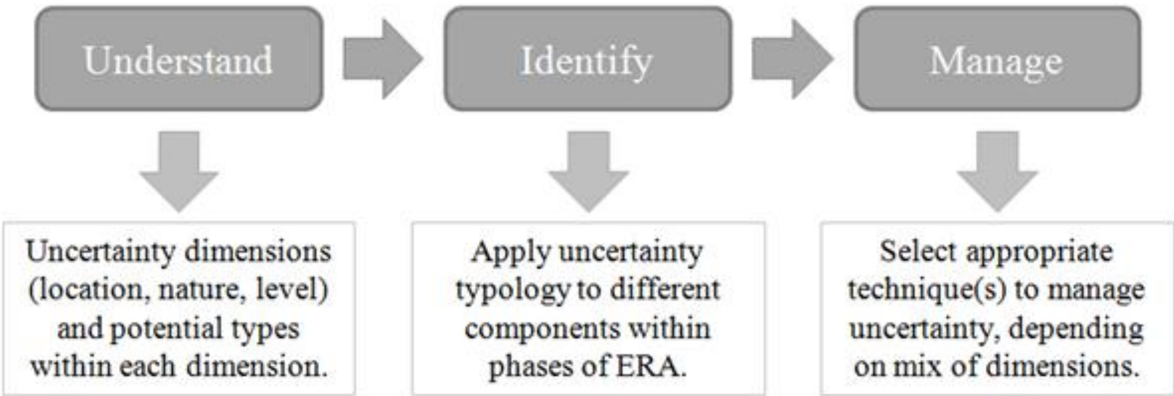


Figure 3.1 A basic overview of uncertainty analysis, including the stages of understanding, identifying, and managing uncertainty, as performed in environmental risk assessments.

Recent research has identified some of the issues associated with the use of uncertainty typologies, for example, that their creation can rely on potentially subjective expert judgement (Knol *et al.* 2009), that their successful implementation can depend on the skill and experience of the end-user (Gillund *et al.* 2008), and that no typology exists which "includes all of its meanings in a way that is clear, simple, and adequate for each potential use of such a typology" (Petersen 2006). However, the full extent of these problems and their potential impacts are yet to be explored in detail across the existing set of typologies.

This chapter, which builds on the uncertainty-based concepts introduced in Chapter 2, presents a critique of existing uncertainty typologies by systematically reviewing distinct versions – on an individual basis and in tandem with one another – and the uncertainties they communicate across a range of ERA domains, including integrated modelling, human and ecological risk assessment, and policy analysis. In doing so, this chapter: (i) explores the intra- and inter-typology conflicts; (ii) examines their applicability to ERAs (drawing from uncertainties identified within existing assessments, described in Sections 2.5-2.7 of Chapter 2); and (iii) provides suggestions for adding value to uncertainty typologies and therefore uncertainty analysis moving forward.

3.2 Method

Uncertainty typologies were selected (Table 3.1) that were either based in the domain of environmental risk or that make specific reference to it. The typologies were published in peer-reviewed articles or books and were sourced using online academic search engines, including Scopus, Web of Science and Google Scholar. Thorough reference-checking of all sources was performed, which ensured the identification of relevant typologies not uncovered by the initial searches. Some sources are not labelled as typologies by their authors, but as uncertainty-based guidance relating to a particular domain, whilst others are presented as original typologies that are explained and justified in full. Each typology was analysed for the type of uncertainties it contained and the frequency of uncertainty dimensions communicated, both in the context of their intended domain (e.g. policy analysis) and their applicability to ERAs.

Table 3.1 Information about the 30 uncertainty typologies examined in this chapter, including the risk domains in which they are based, the methods for data sourcing, and the uncertainties that they contain**.**

Typology source	Risk research domain	Typology based on:	Uncertainties considered
Vesely and Rasmuson 1984	Environmental risk management	Empirical evidence (literature)	Data; Model (understanding, approximation); Completeness; Physical variability
Henrion and Fischoff 1986	Uncertainty analysis	Empirical evidence (literature)	Random; Systematic
Alcamo and Bartnicki 1987	Environmental modelling	Researcher opinion	Model (structure, parameters, forcing, initial state, operation)
Beck 1987	Environmental modelling	Empirical evidence (literature)	Model (aggregation, structure, numerical, parameter); Variability; Errors;
Morgan and Henrion 1990	Environmental policy analysis	Researcher opinion	Statistical variation; Systematic error; Linguistic; Variability; Inherent randomness; Disagreement; Model (approximation, form)
Finkel 1990	Environmental risk management	Empirical evidence (literature)	Model; Parameter; Decision; Natural variability
Funtowicz and Ravetz 1990	Environmental policy analysis	Researcher opinion	Inexactness; Unreliability; Border with ignorance
Wynne 1992	Environmental policy analysis	Researcher opinion; Empirical evidence (literature)	Risk; Uncertainty; Ignorance; Indeterminacy;
Helton 1994	Uncertainty analysis	Researcher opinion	Stochastic; Subjective

Hoffman and Hammonds 1994	Uncertainty analysis	Researcher opinion	Lack of knowledge; Variability
Rowe 1994	Uncertainty analysis	Researcher opinion	Temporal; Structural; Metrical; Translational
Faucheux and Froger 1995	Environmental decision making	Empirical evidence (literature)	Ignorance; Strong uncertainty; Uncertainty; Certainty
van der Sluijs 1997	Environmental modelling	Researcher opinion; Empirical evidence (existing typologies)	Inexactness; Unreliability; Ignorance; Model (input data, conceptual model structure, technical model structure, bugs, model completeness)
Stirling 1998	Environmental policy analysis	Researcher opinion; Empirical evidence (literature)	Risk; Uncertainty; Ambiguity; Ignorance
Bedford and Cooke 2001	Environmental risk management	Researcher opinion	Aleatory; Epistemic; Parameter; Data; Model; Ambiguity; Volitional
Huijbregts *et al.* 2001	Environmental risk management	Researcher opinion	Parameter; Model; Choices; Variability (spatial, temporal, between source and object)
Bevington and Robinson 2002	Uncertainty analysis	Researcher opinion	Systematic errors; Random errors
Regan *et al.* 2002	Environmental risk management	Researcher opinion; Empirical evidence (literature)	Epistemic (measurement error, systematic error, natural variation, inherent randomness, model, subjective judgement); Linguistic (vagueness, context dependence, ambiguity, underspecificity, indeterminacy of theoretical terms)
van Asselt and Rotmans 2002	Environmental modelling	Empirical evidence (existing typologies)	Variability (nature, cognitive, behavioural, societal, technological); Knowledge (inexactness, lack of measurements, practically immeasurable, conflicting

			evidence, ignorance, indeterminacy)
Janssen *et al.* 2003	Environmental risk management	Researcher opinion; Empirical evidence (literature)	Statistical; Scenario; Recognised ignorance; Knowledge-based; Variability-based; Context; Expert judgement; Model (structure, technical, parameters, input); Data; Outputs
Walker *et al.* 2003	Environmental decision making	Researcher opinion; Empirical evidence (literature)	Statistical; Scenario; Recognised ignorance; Total ignorance; Epistemic; Variability; Context; Model (structure, technical, parameters, input, outputs)
Brown 2004	Uncertainty analysis	Researcher opinion; Empirical evidence (literature)	Bounded uncertainty; Unbounded uncertainty; Indeterminacy; Ignorance
Dewulf *et al.* 2005	Environmental risk management	Empirical evidence (existing typologies)	Inherent nature of phenomena; Lack of knowledge; Ambiguity in system understanding
Beer 2006	Environmental risk assessment	Empirical evidence (literature, existing typologies)	Probabilistic; Ambiguity; Incertitude; Ignorance; Indeterminacy
Petersen 2006	Environmental modelling	Researcher opinion; Empirical evidence (literature, existing typologies)	Location; Nature; Range; Recognised ignorance; Methodological unreliability; Value diversity
Hayes *et al.* 2006	Environmental risk assessment	Researcher opinion; Empirical evidence (literature)	Linguistic; Variability; Incertitude
Maier *et al.* 2008	Environmental decision making	Researcher opinion; Empirical evidence (literature)	Data (measurement error, type of data, length of record, analysis); Model (method, record quality, calibration, validation, experience); Human (stakeholder, politics)

Ascough II *et al.* 2008	Environmental decision making	Empirical evidence (existing typologies);	Knowledge; Variability; Linguistic; Process; Model; Variability; Linguistic; Decision
Brouwer and Blois 2008	Environmental modelling	Empirical evidence (existing typologies);	Statistical; Scenario; Qualitative; Recognised ignorance
Knol *et al.* 2009a	Environmental risk assessment	Researcher opinion; Empirical evidence (existing typologies)	Statistical; Scenario; Recognised ignorance; Epistemic; Ontic (process, normative); Model (structure, parameters, input data); Methodological; Analyst uncertainty

3.3 Analysis of existing uncertainty typologies

3.3.1 Comparison of uncertainty terms used

When there is contradiction between different characterisations of uncertainty it can result in confusion and in the use of inappropriate definitions, potentially leading to the selection of inappropriate uncertainty management tools. Contradictions either exist where one term is used for a range of different uncertainties (the same term has multiple definitions; Table 3.2), or where several distinct terms are used to describe the same uncertainty (different terms have the same definition; Table 3.3). Discrepancies between the terminology used in the 30 uncertainty typologies are noted in this section.

Table 3.2 Contradictions in terms (where the same single term is used to represent two or more distinct uncertainty types) between the 30 typologies, with regard to the location, level, and nature of uncertainty, and featuring the terms ambiguity (as denoted by the symbol ①), indeterminacy (②), natural variability (③), parameter (④), random (⑤), statistical (⑥), systematic (⑦), and variability (⑧).

Uncertainty type / Typology	Location							Level					Nature	
	System	Data	Model	Human	Language	Variability	Decision	State 1 (determinism)	State 2 (statistical)	State 3 (scenario)	State 4 (recognised ignorance)	State 5 (total ignoranc	Knowledge-based	Randomness-based
Vesely and Rasmuson 1984														⑧
Alcamo and Bartnicki 1987			④											
Beck 1987			④											
Morgan and Henrion 1990		⑥⑦				⑧								
Finkel 1990		⑥⑦	④			③								
Wynne 1992												②		
Hoffman and Hammonds 1994														⑧
Rowe 1994						③								
Stirling 1999									①					

Bedford and Cooke 2001			④		①									
Huijbregts *et al.* 2001			④											
Bevington and Robinson 2002													⑦	⑤
Regan *et al.* 2002		⑦			①②	③								⑤
van Asselt and Rotmans 2002						③⑧						②		⑧
Janssen *et al.* 2003			④						⑥					⑧
Walker *et al.* 2003			④						⑥					⑧
Brown 2004												②		
Dewulf *et al.* 2005	①													
Petersen 2006									⑥					
Hayes *et al.* 2006														⑧
Ascough II *et al.* 2008			④		①	③⑧								
Brouwer and Blois 2008								⑥						
Knol *et al.* 2009a			④					⑥						

Table 3.3 Identified contradictions in description terms between the 30 typologies where distinct terms are used to describe the same uncertainty type, with regard to the nature dimension (knowledge or randomness) of uncertainty.

Typology	Term describing knowledge uncertainty	Term describing random uncertainty
Vesely and Rasmuson 1984	Completeness	Variability
Helton 1994	Subjective	Stochastic
Hoffman and Hammonds 1994	Knowledge	Variability
Rowe 1994	Completeness	-
van Asselt and Rotmans 2002	-	Variability
Bedford and Cooke 2001	Epistemic	Aleatory
Bevington and Robinson 2002	Systematic	Random
Regan *et al.* 2002	-	Random
Janssen *et al.* 2003	Knowledge	Variability
Walker *et al.* 2003	Epistemic	Variability
Dewulf *et al.* 2003	Uncertainty	Indeterminacy
Petersen 2006	Epistemic	Ontic
Hayes 2006	Incertitude	Variability
Ascough II *et al.* 2008	Epistemic	Aleatory
Knol *et al.* 2009a	Epistemic	Ontic

With reference to the severity of the uncertainty, the term statistical is used to represent both the state of determinism (State 1; Brouwer and Blois 2008; Knol *et al.* 2009a) and the state in which probabilities can be defined but outcomes remain unclear (State 2; Walker *et al.* 2003; Janssen *et al.* 2003; Petersen 2006; Table 3.2). Similarly, the term ambiguity is used to refer to the same latter state (State 2; Stirling 1999), and also the state in which outcomes can be defined but associated probabilities remain unresolved (State 3; Beer 2006). Furthermore, Beer (2006) makes use of the term incertitude to describe a single level of severity (State 2), while Stirling (1999) uses the same term to describe the uncertainties across all levels. Contradictions among typologies based in different research domains might be expected, since they will invariably involve different processes and concerns. However, the severity of uncertainty is a concept adopted across the risk landscape, far beyond the boundaries of environmental concern. When two related typologies, based in risk analysis (Stirling 1999) and ecological risk assessment (Beer 2006) respectively, fail to agree on simple terms, one may assume that these issues transcend environmental risk-based systems.

The term ambiguity is also used when describing the location in which the uncertainties manifest, with specific reference to a system-related uncertainty (Dewulf *et al.* 2005) and a language-related uncertainty (Bedford and Cooke 2001; Regan *et al.* 2002; Ascough II *et al.* 2008). This term carries four separate meanings across the six typologies in which it features (Table 3.2). Similarly, the term statistical is further employed to represent a form of data uncertainty (Morgan *et al.* 1990; Finkel 1990), resulting in three different definitions across seven typologies. Alternate interpretations are also presented for the terms indeterminacy, random, variability, and systematic, with very different meanings attached to each. For example, systematic has been used to either refer to a form of data uncertainty (Henrion and Fischoff 1986; Morgan *et al.* 1990; Finkel 1990; Regan *et al.* 2002), or to the epistemological nature of the uncertainty (Bevington and Robinson 2002).

There are commonalities between the typologies. For example, parameter uncertainty is listed by nine typologies, with all nine agreeing on its use. The uncertainty relating to the inherent variability of natural systems also has a single associated term, natural variability, which is adopted by five typologies. However, the use of competing terms to describe the same uncertainties is commonplace. In fact, of all of the uncertainties communicated in the presented typologies, parameter and natural variability are the only two terms which are consistent across the typologies in which they appear. Epistemic, used to describe knowledge-based failings is used in five separate cases, although, six competing terms are presented by another eight typologies, resulting in the use of seven terms over 13 typologies to describe the same type of uncertainty (Table 3.3). A similar pattern is observed for the terms associated with the aleatory uncertainty bracket, with six separate terms used across 14 typologies. Of the 188 (non-distinct) uncertainties communicated by the 30 typologies, 98 (52%) are contained within just eight representations (Rowe 1994; van Asselt and Rotmans 2002; Walker *et al.* 2003; Janssen *et al.* 2003; Petersen 2006; Maier *et al.* 2008; Ascough II *et al.* 2008; Knol *et al.* 2009a). Within these eight typologies, only 14 terms (with 38 separate occurrences out of the 98) are used to describe the same form of uncertainty in more than one typology. Therefore, there are 60 instances in which either different terms are used to refer to the same uncertainty, or the same terms are used to refer to different uncertainties. In either case, the terminology in these eight typologies disagrees 61% of the time.

Confusion in terms, and possibly in their associated definitions, could easily lead to the adoption of incorrect techniques for managing (i.e. to characterise and possibly reduce) the uncertainties in question, and an unsubstantiated level of confidence being attributed to their quantification and/or reduction.

3.3.2 Comparison of uncertainty frequencies communicated

The existence of multiple typologies across closely related fields presents further challenges. A typology should communicate all uncertainties that are relevant to the domain in which it is based, or be explicit in its limitations, thereby reducing the possibility that it is perceived as comprehensive. Different domains can have distinct methods, processes, concerns, and ultimately different uncertainties. To that end, one might expect agreement between typologies that belong to the same or extremely similar fields. However, this research has identified that this is rarely the case (see Section 3.3.1). For example, typologies that focus on computational modelling procedures show considerable variation in the number of uncertainty categories listed, ranging from four (Brouwer and Blois 2008) to 13 (Walker *et al.* 2003). Within this same domain, variation is also seen in the types of uncertainty considered: Maier *et al.* (2008) describe 11 types of uncertainty, all related to the location in which they manifest, while Petersen (2006) documents eight, which relate to location, severity, and nature. If one of these typologies can be considered comprehensive, the other must either be incomplete or excessive. However, the frequency of uncertainty categories included is not the only metric by which the completeness of a typology should be judged.

3.3.3 Comparison of information sourcing techniques

Each of the 30 identified typologies that were investigated used a limited body of evidence which was sourced in three ways (Table 3.1). Most commonly, the views and opinions of a relatively small number of researchers were used to develop the typology. These typologies can potentially suffer several uncertainties themselves, including subjectivity, intentional bias, and the ability of researcher(s) to communicate effectively. Secondly, small-scale literature reviews (compared to the available relevant body of evidence) were conducted across a relatively restricted topic. In such cases, inappropriate materials selection or a lack of quality sources may result in an incomplete typology. Furthermore, the content may only be applicable to the subject domain of the literature (e.g. atmospheric modelling, habitat conservation, toxicology assessments), and may reduce the accuracy when applying the typology to other domains. Finally, existing typologies were combined from related but non-identical research domains. Arguably the most comprehensive and robust typology presented here (Ascough II *et al.* 2008, on the basis of the defined criteria) is sourced in this way. Whilst this ensures that a large body of relevant research is taken into consideration, the reliability of the output relies on the accuracy of the input. In this sense, combining existing information can mean that shortcomings are transferred into the new typology which may impact on its use.

3.3.4 Suitability of existing uncertainty typologies for environmental risk assessments

The uncertainties discussed in Chapter 2 were shown, through the use of simple examples, to exist in peer-reviewed ERAs (see Sections 2.5-2.7). This sub-section explores the ability of the typologies reviewed in the current chapter to account for such uncertainties.

The identified typologies largely focus on specific aspects of the environmental management process, such as modelling, decision-making, or policy setting. As a result, those typologies are not comprehensive in the context of ERAs, since many processes (and by extension uncertainties) are not covered. Ten of the typologies reviewed here are directly based in either ERA or environmental risk management (incorporating ERA). Of these ten, three provide extensive descriptions of potential uncertainties across all three dimensions. However, of these three, two do not communicate language uncertainties (Janssen *et al.* 2003; Knol *et al.* 2009), and the other does not include a comprehensive description of modelling uncertainties (Regan *et al.* 2002). Therefore, none of the 30 individual typologies analysed in the chapter depicts the full range of potential uncertainties within ERAs. There is certainly scope for creating an improved uncertainty typology.

3.4 Potential improvements to uncertainty typologies

3.4.1 Using the evidence base

Basing the content of typologies on researcher views, small-scale literature reviews, or existing typologies has implicit problems, as previously discussed. There exists a large evidence base of peer-reviewed environmental risk-based research. A structured interrogation of this evidence base, which spans a number of approaches (e.g. modelling, assessments, and management) and environmental concerns, would enable a more comprehensive characterisation of potential uncertainties. Specifically, the analysis of information from competing (but related) research domains will ensure that the full scope of uncertainties are identified. Additionally, such a typology would be able to point to individual occurrences within the evidence base, making it defensible and transparent.

3.4.2 Incorporating factors that influence uncertainty

Extending the traditional typology format to include a system for direct identification of uncertainties would help to minimise any intentional or unintentional bias on the part of the analyst. By analysing the evidence base for any relationships that exist between identified uncertainties and other aspects (e.g. sources, pathways, receptors, the evidence utilised), key associations can be established and statistically evaluated through bivariate analysis. Strong relationships, if deemed to be transferrable, may then form the basis of such an identification system.

3.4.3 Structuring uncertainty typologies

Uncertainty identification requires a level of subjectivity on the part of the practitioner – even when typologies are adopted – and can be further influenced by a lack of familiarity with concepts (Gillund *et al.* 2008). Structuring the typology for the risk domain in which it is intended to be used (e.g. ERA) may prove beneficial. Relating the unfamiliar abstract concepts of uncertainty to more familiar processes (e.g. within problem formulation, exposure/effects assessment, and risk characterisation) may make the typology more intuitive to analysts, making it more robust and ultimately more useful during uncertainty identification.

3.5 Conclusion

Uncertainty typologies should be well-defined, defensible, and, most importantly, accurate. This chapter reviews and analyses these tools, which are integral to the identification of uncertainty within environmental risk systems. It has shown that existing uncertainty typologies across environmental risk domains:

(i) use terminology that is often contradictory;
(ii) communicate varying frequencies and dimensions of uncertainties;
(iii) source information from limited data sets; and
(iv) cannot be applied, on an individual basis, to ERAs in order to characterise the wide range of potential uncertainties.

This chapter has highlighted the limitations of existing uncertainty typologies, which is of benefit to environmental risk analysts in their attempts to better qualify uncertainties, and thus statements about risk, within ERAs. However, more could be done to improve uncertainty typologies and the guidance related to their implementation, in the context of ERAs and uncertainty analysis in general.

[This page is intentionally left blank]

Chapter 4: A novel uncertainty typology for environmental risk assessments

4.1 Introduction

An important preliminary stage of the uncertainty management process is the identification of different types of uncertainty (Morgan *et al.* 1990), where it is essential that environmental risk practitioners are able to draw from a clear and defensible typology of uncertainties (Knol *et al.* 2009a). However, as demonstrated in Chapter 3, existing typologies suffer many problems, relating primarily to research domain transferability and content reliability. The primary objective of this chapter is to provide an evidence-based typology of uncertainties across a specific form of ERA, namely weight of evidence assessments (see Section 2.3.2), and in so doing resolve the issues surrounding existing categorisations.

Environmental uncertainty comprises three dimensions, namely the location, nature, and level (Walker *et al.* 2003; Janssen *et al.* 2003; Knol *et al.* 2009a), and different uncertainties must be managed in different ways and with different techniques (van der Sluijs *et al.* 2004; Refsgaard *et al.* 2007; see Section 2.8). Identifying the different types of uncertainties that exist in applied situations is therefore an essential part of the uncertainty management process (Morgan *et al.* 1990). It is the role of the uncertainty typology to aid this process by providing comprehensive, relevant, and reliable categorisations (complete with definitions) of all potential types of uncertainty that may be encountered (van Asselt and Rotmans 2002; Knol *et al.* 2009a; see Sections 2.5-2.7). However, the distinct typologies that exist within the domain of environmental risk-based research (see Table 3.1) suffer many problems (see Section 3.3). If doubt surrounds the legitimacy of the adopted categorisation(s), there must also be concerns about the reliability of the following uncertainty identification.

The evidence-based approach used in this chapter ensures that any assertions made are supported by a defensible set of research articles, and, due to the inclusion of environmental WOE assessments, that these assertions will span a diverse set of interests, making the resulting typology relevant across a number of distinct research domains. Further to the typology, a secondary aim is to provide an analysis of the adoption of uncertainty management techniques (UMTs) used when faced with different uncertainties, in order to highlight any recurring areas of weakness. Finally, this chapter aims to provide a statistical description of the relationships that exist between identified uncertainties and other aspects of the WOE assessments, and to discuss the significance of these associations in an uncertainty identification context.

4.2 Method

4.2.1 Parameters of the evidence base

In order to categorise uncertainty in environmental WOE frameworks, analyse the use of techniques in their management, and describe the relationships between identified uncertainties and other aspects of the assessments, an evidence base of peer-reviewed literature was established. Searches were conducted for directly labelled WOE literature, using online academic databases, using the keywords *weight*, *evidence*, *risk*, and *uncertainty*. Non-labelled WOE literature was also searched for, using the keywords *risk*, *assessment*, and *uncertainty*. In-built filtering within online databases was used to remove obviously non-relevant literature before the remaining articles were assessed for inclusion based on the following criteria:

- the article must include (or be in its entirety) a WOE assessment, consisting of either a qualitative, semi-quantitative, or quantitative methodology (see Section 2.3.2);
- the assessment must make direct reference to all uncertainties to be recorded within this research, thereby minimising researcher-subjectivity when creating the typology;
- the assessment must be original research and not a review of previously published work, in order to avoid duplicate values; and
- an aspect of the environment must feature in at least one part of the S-P-R paradigm (see Section 2.3.1), where the environment "consists of all, or any, of the following media, namely the air, water, or land" (Environmental Protection Act 1990).

These criteria ensured that only original environmentally-focused WOE assessments that specifically mentioned uncertainty were included within this chapter. The general search terms used were chosen to include a wide range of research domains.

4.2.2 *Data collection*

The articles (conforming to the selection criteria) were examined in full and relevant information was extracted and recorded in separate spreadsheet entries (Table 4.1).

Table 4.1 Information extracted from the evidence base of 171 environmental WOE articles, with associated value types of string (text) or Boolean (yes/no).

Information extracted	Description of information	Value type
Uncertainty	Identified uncertainty type	String
UMT	Uncertainty management technique associated with the identified uncertainty type, if any	String
LOE (all)	All lines of evidence employed in the assessment	String
LOE (uncertainty)	All lines of evidence employed in the assessment that are associated with the identified uncertainty	String
Source	Source(s) of (potential) harm	String
Source (uncertainty)	Is the source associated with the uncertainty type?	Boolean
Pathway	Pathway(s) between source and receptor	String
Pathway (uncertainty)	Is the pathway associated with the uncertainty type?	Boolean
Receptor	Recipient(s) of (potential) harm	String
Receptor (uncertainty)	Is the receptor associated with the uncertainty type?	Boolean
WOE	Type of WOE framework used in assessment	String

A working list of definitions was kept to ensure that observations were consistent and distinctions between uncovered uncertainties were not blurred. In evaluating potential relationships between uncertainties and the UMTs, the lines of evidence utilised, and members of the S-P-R paradigm, respectively, a relationship was deemed to exist if the hypothetical removal of the attribute (e.g. an ingestion pathway) from the assessment led to the disappearance of the uncertainty, and not otherwise. For example, to determine if an identified uncertainty in an ecotoxicity assessment was directly linked to an exposure pathway attribute (e.g. ingestion of contaminated materials), the pathway would be hypothetically eliminated from the assessment. If the uncertainty could feasibly be considered to remain, no relationship between the uncertainty and that attribute could be recorded. Importantly, no upper limit was set as to the number of UMTs, LOEs, or members of the S-P-R paradigm that could be associated with any singularly identified uncertainty type.

4.2.3 Data organisation

The uncertainty and LOE data were organised separately using an iterative category clustering technique (Hartigan 1975); no other data required clustering. The different objects (i.e. the uncertainties and the LOEs) were categorised into distinct groups, such that the degree of association between any two objects was maximal if they belonged to the same group and minimal otherwise. In this way, the existing organisation by source article (from the data collection stage) was translated into organisation by relevance to other similar data values. Clustering provides a defensible strategy for organising data. However, clustering of qualitative data (without the use of optimised algorithms) can introduce some user subjectivity. To reduce the potential for subjectivity in assigning objects to groups, the process was performed iteratively, with definitions and categorisations continually refined, rather than in one round of analysis. Clustering of uncertainty data directly enabled the creation of the uncertainty typology.

4.2.4 Data analysis

The frequencies with which the different categories and sub-categories of uncertainties were associated separately with the UMTs, LOEs, S-P-Rs and WOE types were recorded and converted to percentage values of total occurrences in order to identify the most commonly occurring relationships. A separate bivariate analysis was performed using SPSS v19 (SPSS Inc., Chicago IL) to quantify the relationships between all two-variable combinations (α = 0.05).

4.3 Results

4.3.1 Data frequencies and organisation

Uncertainty typology

Analysis of the collected WOE literature (n=171 articles; see Appendix A for the full list of articles), in conjunction with iterative clustering of the extracted data (Figure 4.1), revealed 20 separate types of uncertainty (Table 4.2).

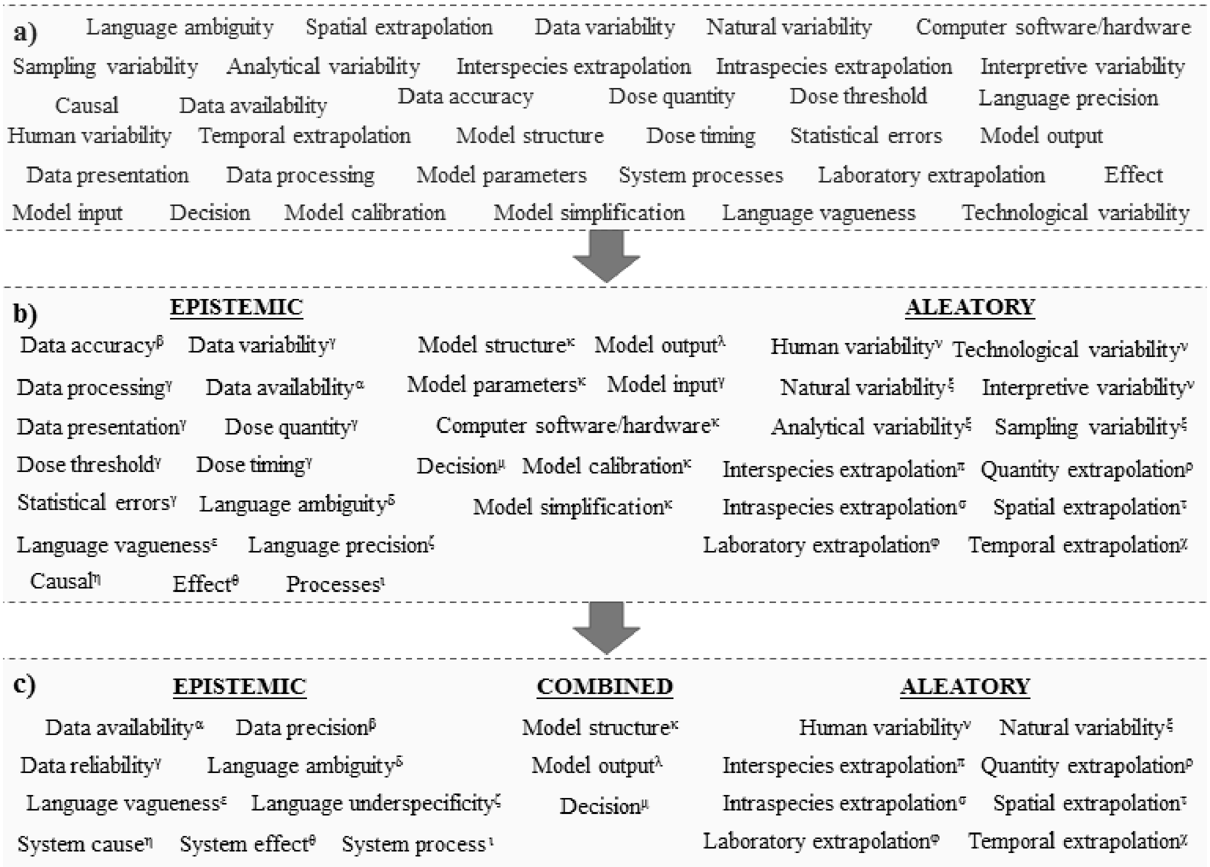


Figure 4.1 Overview of the clustering process applied to the uncertainty data extracted from the collected evidence base (n=171 environmental WOE assessments), showing: a) all 35 recorded location-based uncertainty types; b) all 35 recorded location-based uncertainty types organised according to their nature; and c) final 20 location-based uncertainty types organised according to their nature. The superscript Greek letters in b) are matched to the superscript Greek letters in c), representing clustering into like groups. For example, model structure, model parameters, computer software/hardware, model calibration, and model simplification uncertainties, denoted by the Greek letter Kappa (κ), in b) are clustered into model structure uncertainty, also denoted by κ, in c).

Table 4.2 Typology of uncertainties that exist within the evidence base of 171 environmental WOE articles, with basic definitions.

Nature	Category	Sub-category	Definition
Epistemic	Data	Availability	referring to the incompleteness, scarcity, or absence of data
		Precision	concerning the lack of accuracy or precision in obtained data
		Reliability	reflecting its trustworthiness i.e. data are erroneous for some specified reason
	Language	Ambiguity	where multiple meanings are possible
		Underspecificity	where meanings are not exact
		Vagueness	where meanings are not clear and understandable
	System	Cause	concerning a lack of clarity regarding the source(s) of harm
		Effect	relating to the influence a particular stressor (source) has upon the receptor(s)
		Process	where the risks are not understood or a process vital to a successful assessment is not identified
Aleatory	Variability	Human	which exists through intentionally biased and subjective human actions
		Natural	which pertains to the stochastic traits of natural systems
	Extrapolation	Intraspecies	where information specific to members of a species is used to represent other members of the same species
		Interspecies	where information specific to members of a species is used to represent members of a different species
		Laboratory	where information specific to laboratory conditions is used to represent real-world scenarios
		Quantity	where information specific to one quantity is used to represent another
		Spatial	where information specific to one spatial scale is used to represent another
		Temporal	where information specific to one timescale is used to represent another
Combined	Model	Structure	concerning the representation of real-world processes in model form
		Output	reflecting the level of confidence in the produced results
	Decision	Decision	where doubt surrounds an optimal course of action, often in the face of differing objectives.

These 20 uncertainties had a total of 385 individual occurrences (Figure 4.2). The data uncertainty (n=125 out of 385; 32.5%) and extrapolation uncertainty (n=110; 28.6%) categories were the most frequently occurring, with the decision uncertainty (n=6; 1.6%) and language uncertainty (n=16; 4.2%) categories the least frequent.

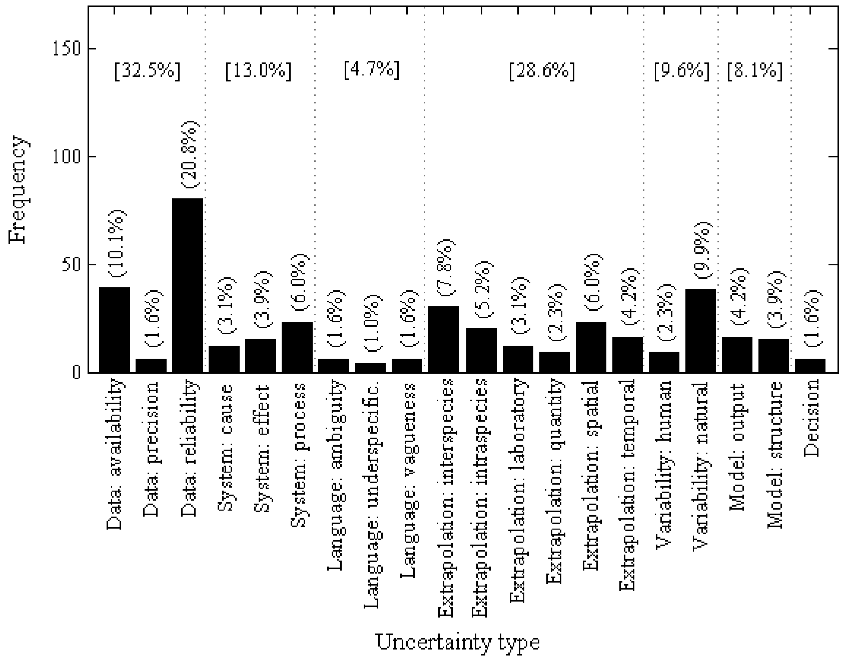


Figure 4.2 Occurrence frequencies of the individual uncertainty types (n=20) within the collected evidence base (n=171 environmental WOE articles), with occurrence proportions of individual uncertainties in parentheses (‘()’) and categories of uncertainties in square brackets (‘[]’).

Uncertainty management techniques

Data extracted from the sources highlighted the use of a variety of UMTs (n=27), with a total of 453 separate applications. Total occurrence frequencies of the mechanisms are shown in Figure 4.3.

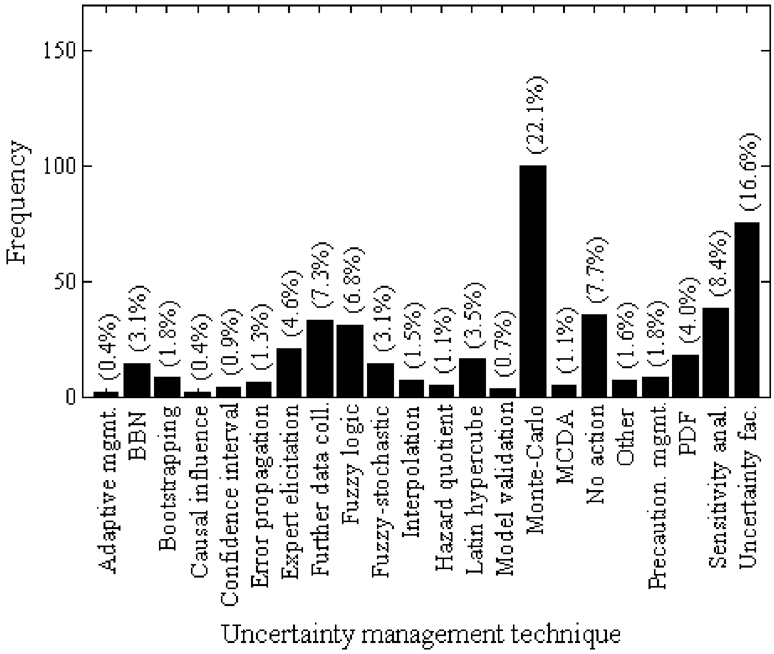


Figure 4.3 Occurrence frequencies of the uncertainty management techniques (n=22) employed within the collected evidence base (n=171 environmental WOE articles), with occurrence proportions in parentheses (‘()’).

MCS was adopted most frequently (n=100 out of 453; 22.1%), followed by uncertainty factors (n=75; 16.6%), sensitivity analysis (n=38; 8.4%), and taking no action (n=35; 7.7%). Brief descriptions and associated uncertainties of the most frequently occurring UMTs are outlined in Table 4.3.

Table 4.3 Descriptions of the most frequently occurring uncertainty management techniques within the evidence base of 171 environmental WOE articles, along with their associated uncertainties. See Section 2.8 for more detailed descriptions.

Uncertainty management technique	Description	Associated uncertainties	Referenced in:
Monte-Carlo simulation	Utilises repeated executions of numerical models to simulate stochastic processes.	Data, Extrapolation, Variability, Model, System	Ma 2002 Qin and Huang 2009
Uncertainty factor	Attaches a factor-based correction to the data being used which reflects the level of uncertainty within it.	Extrapolation, System, Data, Variability	Calabrese *et al.* 1997 Phillips *et al.* 2008
Sensitivity analysis	Tests the sensitivity of a chosen output variable to variations in quantities relating to input variables.	Data, Model, Extrapolation, System	Huysmans *et al.* 2006 Oughton *et al.* 2008
No action	Not attempting to quantify, reduce, or manage uncertainties, whether recognised by the publication author(s) or identified through the research in this chapter.	Data, Extrapolation, System, Variability, Model	Cesar *et al.* 2009
Further data collection	The collection of increased quantities of data.	Extrapolation, Data, Variability	Avagliano and Parella 2009
Fuzzy logic	A form of multi-valued logic that allows its components to be approximate rather than precise.	Data, Language, Model, Variability	Zadeh 1965 Acosta *et al.* 2010
Expert elicitation	Seeks to capture the knowledge of one or more experts in a field with regard to a specific matter.	Data, System , Variability	Kandlikar *et al.* 2007

Probability density function	Describes the frequency of occurrence for different parameter values over a given range.	Data, Variability	Oughton *et al.* 2008
Latin hypercube sampling	Splits a distribution into distinct intervals for sampling and use as inputs to a numerical model.	Data, Variability	Klier *et al.* 2008 Kumar *et al.* 2009
Bayesian belief network	A graphical representation of a system, in which relationships between uncertain characteristics are expressed through probability values.	Variability, Data, System	Aspinall *et al.* 2003
Fuzzy-stochastic system	A hybrid approach for incorporating epistemic and stochastic uncertainties separately.	Data, Extrapolation, Language	Li *et al.* 2007 Kumar *et al.* 2009
Precautionary management	Management based upon the application of the Precautionary Principle.	Extrapolation, System	Godduhn and Duffy 2003
Multi-criteria decision analysis	Brings together criteria and performance scores to provide a basis for integrating risk and uncertainty levels.	Decision	Linkov *et al.* 2007 Critto *et al.* 2007
Adaptive management	Incorporate the needs of many into an iterative system where differing alternatives and objectives are present.	Decision	Dey *et al.* 2000 Williams *et al.* 2009

Lines of evidence

Clustering of the LOEs produced 15 distinct groups (Figure 4.4), each with varying quantities of sub-categories, with a total of 523 instances. Toxicology (n=125 out of 523; 23.9%), chemistry (n=107; 20.5%), and bioaccumulation-related evidence (n=55; 10.5%) were the most frequently occurring LOEs within the dataset.

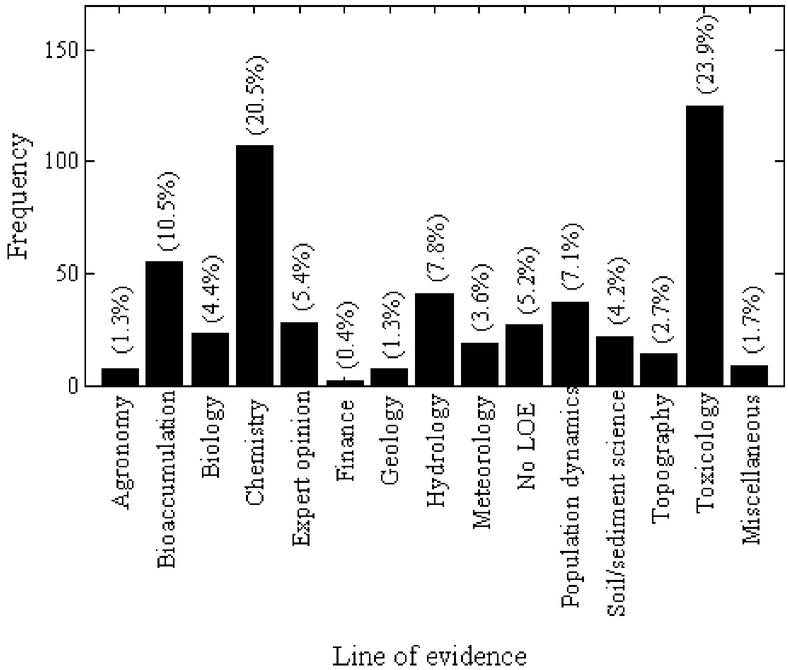


Figure 4.4 Occurrence frequencies of the lines of evidence employed (n=15) within the collected evidence base (n=171 environmental WOE articles), with occurrence proportions in parentheses ('()').

S-P-R paradigm and WOE types

All sources (n=103 out of 374; 27.5%), pathways (n=111; 29.7%), and receptors (n=160; 42.8%) involved in the WOE assessments were recorded and related to each occurrence of the identified uncertainty types. Additionally, the types of WOE frameworks employed were also recorded in relation to all instances of identified uncertainties, and were either qualitative (n=48 out of 385; 12.5%), semi-quantitative (n=138; 35.8%), or quantitative (n=199; 51.7%).

4.3.2 Frequency relationships

Relationships between uncertainties and uncertainty management techniques

The highest frequency relationships between the uncertainty categories and UMTs employed (Figure 4.5) occurred between data uncertainties and MCS (n=56 out of 453 relationships), between extrapolation uncertainties and uncertainty factors (n=40), and between extrapolation uncertainties and MCS (n=18). On a proportional basis, the highest dependencies were seen between language uncertainties and fuzzy logic (68.8%; i.e. language uncertainties were managed with fuzzy logic in 68.8% of cases), model uncertainties and sensitivity analysis (35.1%), and data uncertainties and MCS (34.4%). Overall, uncertainties were associated with at least one UMT in 92.3% of cases, and were therefore unmanaged 7.7% of the time.

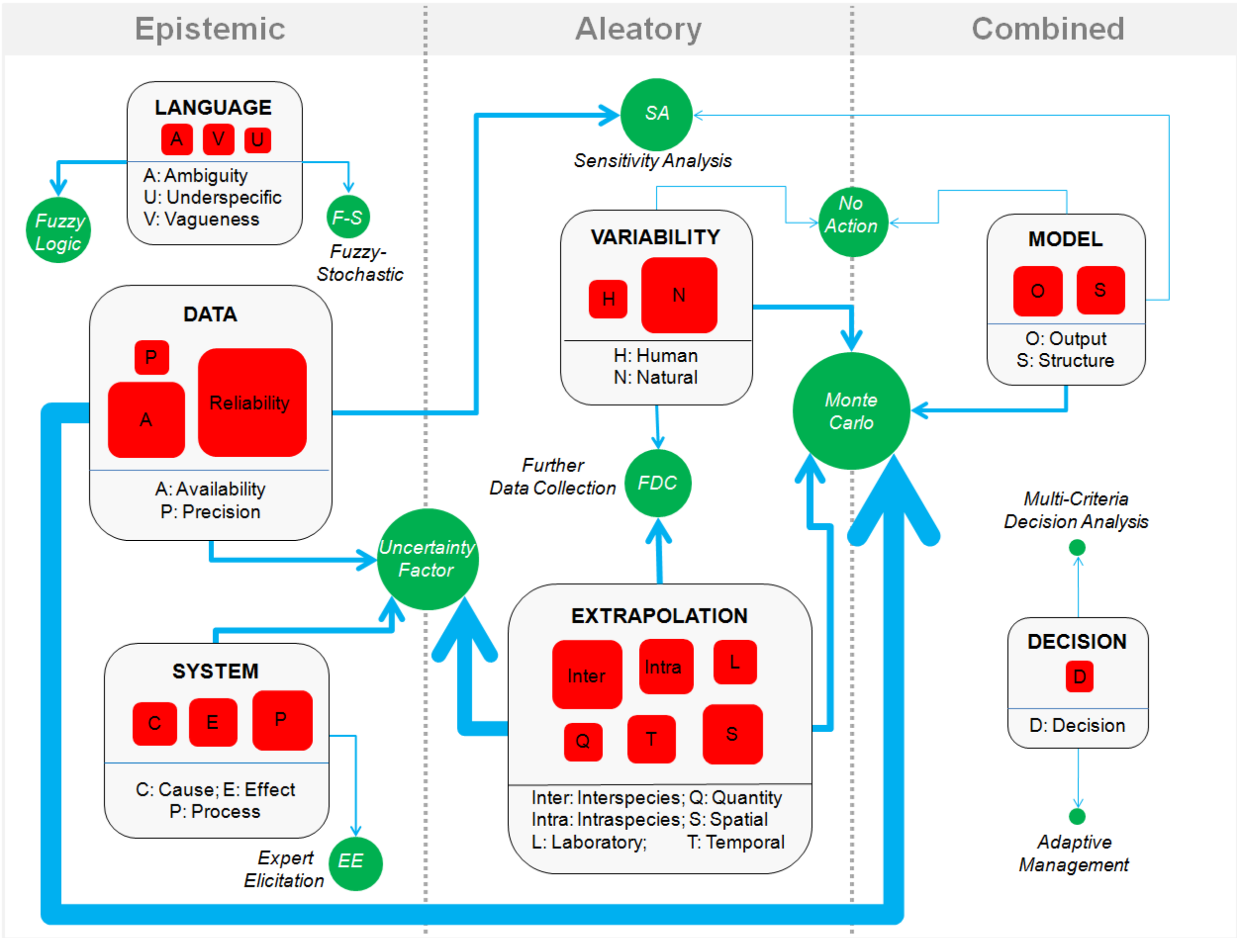


Figure 4.5 The occurrence frequencies of the individual uncertainty types identified (red squares; n=20), management techniques utilised (green circles; n=10), and the relationships between them (blue lines) within the collected evidence base (n=171 environmental WOE articles). The areas of the squares and circles (which depict the respective occurrence frequencies) are relative to each other, as are the widths of the dependency lines, where an increasing (square or circle) area and (line) width indicates an increasing frequency.

Relationships between uncertainties and lines of evidence

The highest frequency relationships between the individual uncertainty types and the LOEs utilised (Figure 4.6) were demonstrated between reliability (data) uncertainty and both chemistry-based evidence (n=42 out of 523 relationships) and toxicological evidence (n=24), and between interspecies (extrapolation) uncertainty and toxicological evidence (n=23). On a proportional basis, the strongest relationships were recorded between human (variability) uncertainty and expert opinion evidence (100%), underspecificity (language) uncertainty and expert opinion (75.0%), and quantity (extrapolation) uncertainty and toxicological evidence (66.7%). Overall, uncertainties were connected with at least one LOE in 93.0% of cases.

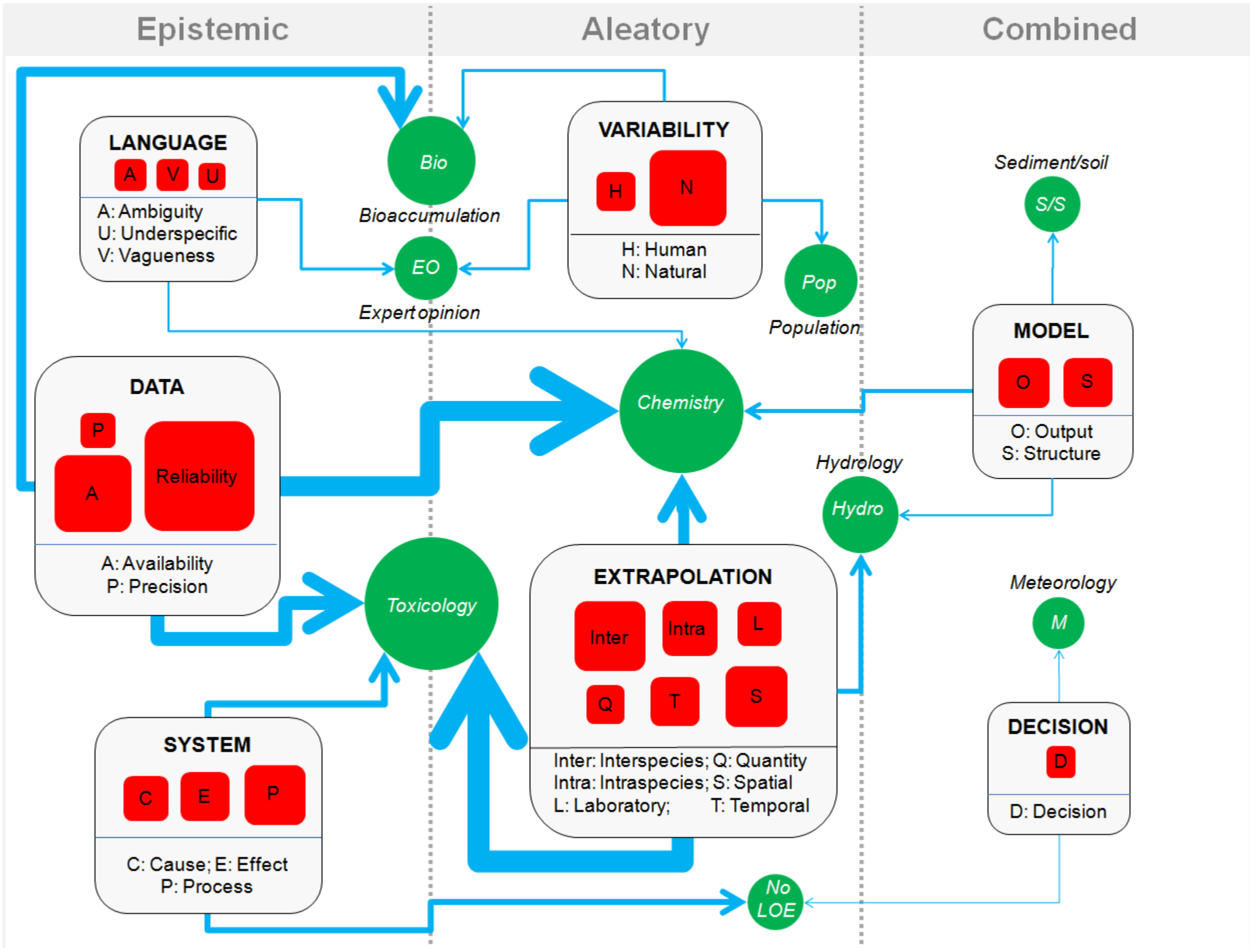


Figure 4.6 The occurrence frequencies of the individual uncertainty types identified (red squares; n=20), lines of evidence utilised (green circles; n=9), and the relationships between them (blue lines) within the collected evidence base (n=171 environmental WOE articles). The areas of the squares and circles (which depict the respective occurrence frequencies) are relative to each other, as are the widths of the dependency lines, where an increasing (square or circle) area and (line) width indicates an increasing frequency.

Relationships between uncertainties and sources, pathways, and receptors

Uncertainties were also shown to be associated with sources (n=103 out of 374 relationships; 26.8%), pathways (n=111; 28.8%), and receptors (n=160; 41.6%). Overall, the identified uncertainties under consideration were associated with at least one aspect of the S-P-R paradigm in 79.5% (n=306 out of 385) of all cases.

4.3.3 Statistical relationships

Relationships between uncertainties and uncertainty management techniques

The strongest correlations between the uncertainty types and UMTs (Figure 4.7) occurred between decision uncertainty and adaptive management (ρ=0.57), spatial (extrapolation) uncertainty and interpolation (ρ=0.46), and cause (system) uncertainty and causal influence (ρ=0.40).

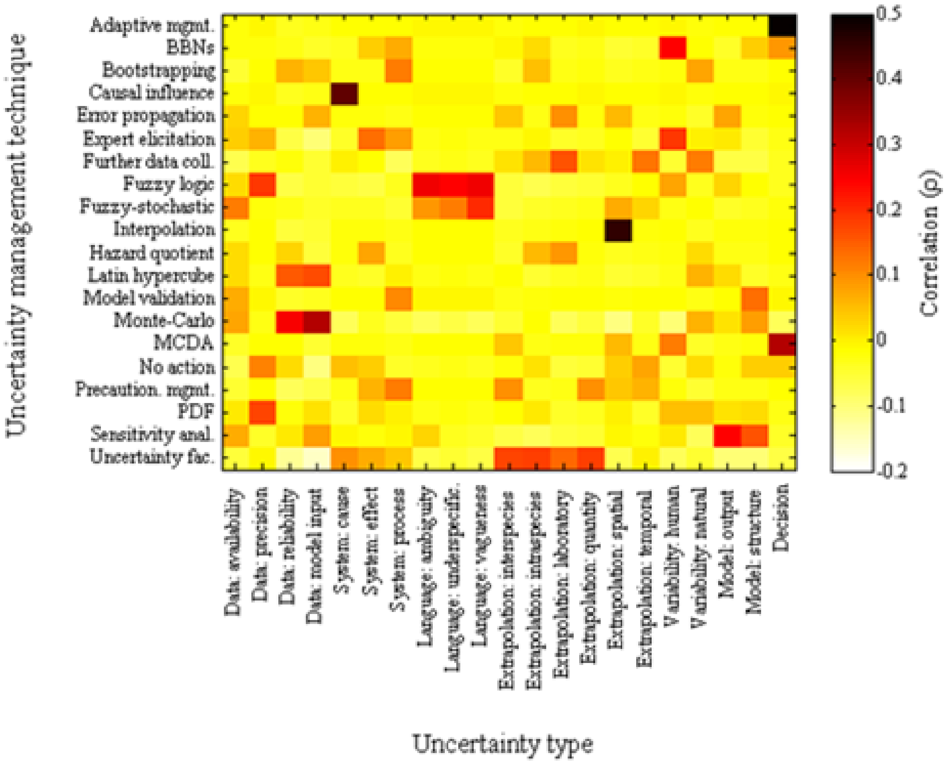


Figure 4.7 Matrix plot showing the correlation values (ρ), where a higher value indicates a stronger correlation, between the uncertainties and their respective management techniques.

A similar strength correlation occurred between the portion of data uncertainties used as parameter values in computational and/or numerical models (and therefore consist of repeated values from within the data category; marked *model input* in Figure 4.7) and MCS (ρ=0.32). Positive correlations were also observed between several uncertainty-category/UMT combinations, where all individual uncertainty types within the category shared a positive correlation with the respective UMT. The strongest of these relationships were language uncertainties with fuzzy logic (ρ=0.45) and fuzzy-stochastic systems (ρ=0.24), and model uncertainty with sensitivity analysis (ρ=0.29).

Relationships between uncertainties and lines of evidence

Analysis of the uncertainty types in concert with the respective LOEs utilised (Figure 4.8) revealed two important correlations, between human (variability) uncertainty and expert opinion (ρ=0.55), and process (system) uncertainty and the 'no LOE' used category (ρ=0.40), respectively.

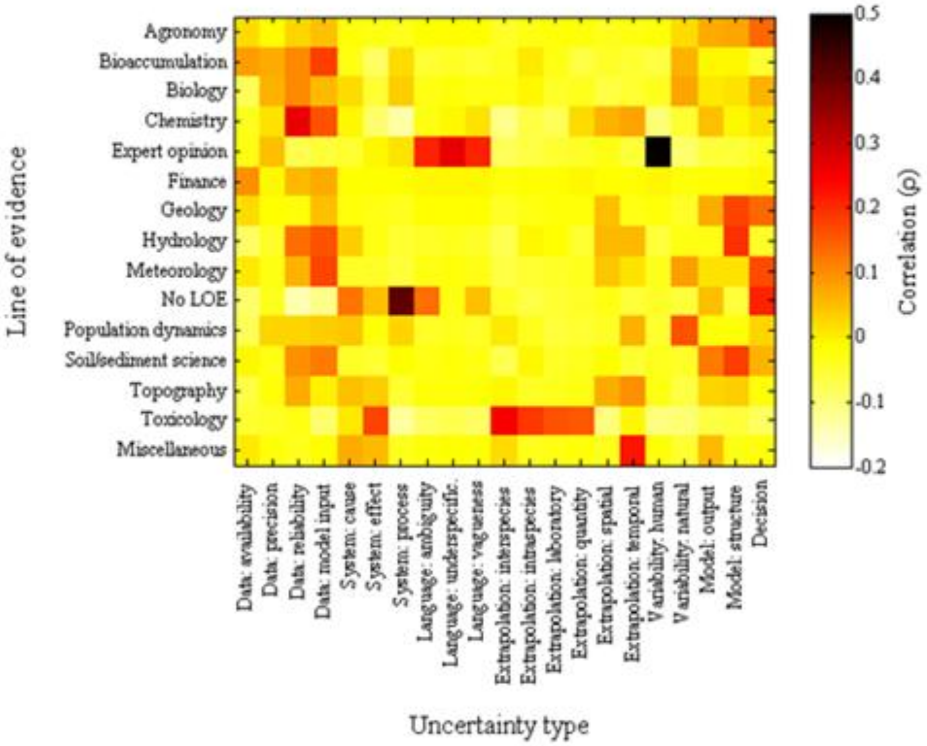


Figure 4.8 Matrix plot showing the correlation values (ρ), where a higher value indicates a stronger correlation, between the uncertainties and their associated lines of evidence.

There were also several positive correlations between uncertainty categories and LOEs, the strongest of which were noted as: system uncertainties with the 'no LOE' used category ($\rho=0.38$); language uncertainties with expert opinion ($\rho=0.39$); and model uncertainties with six separate evidence types (ranging from $\rho=0.01$ to $\rho=0.22$).

Relationships between uncertainties and sources, pathways, receptors, and WOE frameworks

No meaningful correlations were observed between the S-P-R paradigm and individual uncertainty types, and only two relationships of any meaning were witnessed with the uncertainty categories, namely data uncertainty and pathways ($\rho=0.18$) and extrapolation uncertainty and receptors ($\rho=0.20$).

The strongest correlation between uncertainty types and the WOE frameworks employed was between the data used as model input and quantitative frameworks ($\rho=0.31$). The data, language, and model uncertainty categories saw positive correlations with quantitative ($\rho=0.31$), semi-quantitative ($\rho=0.12$), and quantitative ($\rho=0.19$) frameworks, respectively.

4.4 Discussion

4.4.1 The uncertainty categorisations within the developed typology

Whilst many of the uncertainty categories and sub-categories contained in the developed typology are similar to those seen in other existing typologies (see Table 3.1), there are differences. For example, language uncertainties (if included at all) were typically separated into their own category (e.g. Morgan *et al.* 1990; Regan *et al.* 2002; Ascough II *et al.* 2008), but are here deemed to be epistemic. The uncertainties associated with language arise for a number of reasons, but stem primarily from a lack of clarity (Morgan *et al.* 1990). However, the definitions, contexts, and applications associated with language can be controlled (Regan *et al.* 2002). Theoretically, language uncertainties can be quantified, reduced or even removed – techniques such as fuzzy logic are testament to this – equating them with the other uncertainties (data and system) within the epistemic set. Despite their relatively low levels of occurrence within the WOE evidence base (of just 4.7%; Figure 4.2), communicating the epistemic quality of language uncertainties allows analysts to approach them with reduction and elimination in mind, which may previously not have been the case.

The extrapolation sub-category in the developed typology is a suggested divide of the aleatory category, where previously it has been grouped with model uncertainties (Regan *et al.* 2002; Walker *et al.* 2003), treated as a branch of variability (Huijbregts *et al.* 2001), or more commonly ignored altogether. Extrapolation can be considered an attempt at rectifying availability issues: if information were readily available, extrapolation would not be necessary. However, when it is required, the process is deemed uncertain due to the natural variability involved (e.g. spatially and temporally extrapolating meteorological data beyond the physical limits of an existing network of measuring stations to a study site). Extrapolation can therefore be considered the result of epistemic failings, with the connected uncertainties driven through aleatory means. Whilst an increase in relevant epistemic knowledge may prevent the need for extrapolation (thereby providing a distinction from aleatory variability, which can be neither eliminated nor reduced), when it is required it is the aleatory-based failings that must be addressed. These observations confirm extrapolation uncertainties to be aleatory in nature, and indicate that they should be considered separately from the aleatory variability category. The quantity, spatial, and temporal extrapolation uncertainties featured in the uncertainty typology directly correspond to the changing scales of risks (Rowe 1994), a concept that is often overlooked. Extrapolation uncertainties are common (with 28.6% of all occurrences within the WOE evidence base, second only to data uncertainties with 32.5%; Figure 4.2), and so communicating their existence is important. Furthermore, categorising them as aleatory in nature ensures that they can be managed in an appropriate way.

Other notable differences between the locations of uncertainty seen in the typology developed in this chapter and those discussed elsewhere (see Section 2.5) include:

- the introduction of a data reliability sub-category, which accounts for 20.8% of all uncertainties within the WOE evidence base (Figure 4.2), and primarily reflects the measurement and systematic sub-categories seen within existing typologies;
- the identification of the cause, process, and effect sub-categories within system uncertainty, where previously no sub-categories were recognised;
- a lack of epistemic human-based uncertainties, which were only identified as being aleatory in nature within the WOE evidence base;
- a lack of technological and institutional uncertainties within the aleatory variability category, due to none being identified within the WOE evidence base; and

- the inclusion of a model structure sub-category, which primarily reflects the structural and technical sub-categories seen within existing typologies, and the added inclusion of a model output sub-category, which is previously unseen.

The final major distinction comes through the inclusion of a combined epistemic and aleatory category, containing model and decision uncertainties. These uncertainties have the potential to incorporate both epistemic and aleatory aspects, forcing a separation from those sets. For example, modelling may incorporate system uncertainty, which can reduce confidence in the structure of a model, as well as variability uncertainty, which may cast doubt over the validity of the model's output. Reducing secondary uncertainties associated with incorporated groups is therefore just as important as managing the primary failings.

Every identified uncertainty with defined nature and location-type must also be considered in terms of its level (i.e. severity; Walker *et al.* 2003; Janssen *et al.* 2003; Refsgaard *et al.* 2007). The level of an identified uncertainty is highly context-dependant and cannot, at present, be ascribed *a priori*. Owing to this, there is a reduced need (compared with the nature and location) for an uncertainty typology to make specific reference to potential levels. It may simply be more appropriate to do it in an accompanying narrative. However, when the focus shifts from uncertainty identification to uncertainty management, an effective typology should also aim to communicate methods for quantification and/or reduction. Therefore communicating uncertainty levels is essential as a change in level will cause a change in the optimal UMT. In terms of data uncertainties, for example, when there is a level of statistical uncertainty (i.e. where the uncertainty can be adequately described in statistical terms) the associated data uncertainty can be tackled through sensitivity analysis. However, if we were in the range of scenario uncertainty (i.e. where it is not possible to formulate probabilities), scenario analysis, for example, would be more appropriate (Refsgaard *et al.* 2007). Ultimately, selection of a suitable UMT is dependent on the mix of all three uncertainty dimensions: location, nature and level.

4.4.2 A novel approach for characterising uncertainty

The existing uncertainty characterisations (see Table 3.1) are predominantly based within specific research areas, using categorisations that are primarily relevant to those fields. They communicate varying frequencies of uncertainties, often in a contradictory fashion, and use a number of different approaches in their construction, including small-scale literature reviews (e.g. Regan *et al.* 2002) and amalgamations of existing frameworks (e.g. Ascough II *et al.* 2008). This has led to overlapping and contradictory sets of categorisations. The uncertainty typology (Table 4.2), resulting from the analysis of 171 WOE articles and iterative clustering of the subsequent uncertainty dataset, addresses these issues in the following ways:

- Firstly, the developed typology does not restrict observations to narrowly-defined research domains (e.g. conservation biology) but instead extends the focus to all concerns of an environmental nature, enabling the typology to be more transferrable and relevant to a larger number of risk analysts.
- Secondly, using WOE assessments, which contain a variety of assessment approaches and techniques as well as distinct forms of evidence, increases the potential for a larger spectrum of uncertainties to exist. This is reflected in the typology which, containing 20 distinct forms of location-based uncertainties arranged according to their natures, is the most extensive to date.
- Finally, by constructing and interrogating a large supporting evidence base of peer-reviewed articles (n=171) all uncertainty categorisations within the typology are supported and defensible.

4.4.3 The appropriateness of uncertainty management techniques employed

UMTs should be used in concert with specific types of uncertainty (Refsgaard *et al.* 2007). The correct adoption of any one UMT is therefore dependent upon the uncertainties present. The occurrence frequency analysis and statistical analysis conducted between the uncertainty types and UMTs highlighted several relationships, the vast majority of which show the UMTs being used to tackle appropriate uncertainties (see Figures 4.5 and 4.7). This is a positive finding, since the incorrect utilisation of a UMT may be considered just as important as choosing not to use one at all, which was the fourth most-adopted option. In this research, taking 'no action' is defined as the publication author(s) recognising uncertainties but not taking action, with or without offering justification. As well as indicating the inappropriate use of this technique with reference to specific uncertainties (primarily model and variability), the occurrence frequency analysis and resulting dependency diagram (Figure 4.5) convey a more important point: dealing with uncertainties should be a major priority within these assessments. The fact that the 'no action' mechanism appears so often suggests that this is not currently the case.

4.4.4 Separating uncertainty and variability

The categorisation of uncertainties as being either epistemic, aleatory, or a combination of the two, might imply that each of the identified UMTs can equally be assigned to one of these groups. This is not the case, nor is there a single mechanism that offers comprehensive solutions to all of the identified uncertainties.

Whilst uncertainties appear to fall easily into the aforementioned groupings, the boundary can be less well defined in applied situations (Wu and Tsang 2004; Merz and Thieken 2009). The most pertinent example of this is the use of MCS in an attempt to cope with both forms of uncertainty. Since epistemic and aleatory uncertainties can both be described by probability distributions, many assessments involving a first-order Monte-Carlo procedure claim to successfully handle both (Wu and Tsang 2004). However, the ensuing single distribution (which may combine data reliability uncertainty with inherent natural variability) incorrectly implies that uncertainty and variability are the same, and that they can be dealt with as one (Wu and Tsang 2004). Problems may still exist even when a distinction is made: incorrectly treating variability as if it were uncertainty may yield a meaningless distribution when a single figure is required (Vose 2000). Effectively, the techniques that are employed to manage uncertainty can, if executed incorrectly, introduce further errors.

It is increasingly recognised that uncertainty and variability need to be treated separately (Kelly and Campbell 2000; Li *et al.* 2008; Kumar *et al.* 2009; Qin and Huang 2009). Once separated, both aleatory variability and epistemic uncertainty can be quantified, and steps can be taken to reduce and potentially remove epistemic uncertainty. Techniques such as second-order Monte-Carlo (Griffin *et al.* 1999; Wu and Tsang 2004) and integrated fuzzy-stochastic systems (Li *et al.* 2007; Kumar *et al.* 2009; Qin and Huang 2009) have emerged that can manage both aleatory and epistemic uncertainties. Moreover, through correct uncertainty management, they attempt to eliminate the inferred, and potentially unjustifiable, level of confidence that can incorrectly be assigned to risk estimates.

4.4.5 Uncertainty and evidence: moving from characterisation to identification

The observation that uncertainties are more often connected to the LOEs used in the risk assessments (93.0%) than to the S-P-R paradigm (79.5%) or even to the techniques used to handle their presence (92.3%) suggests interrelatedness between uncertainty and evidence. The relationship between uncertainty and evidence employed would be even stronger were it not for the correlation (ρ=0.40; Figure 4.8) between process (system) uncertainties and the 'no LOE' category. It is not surprising that this correlation exists, since being uncertain about a system process is the obvious result of insufficient evidence about the system (and in particular the process) under study. However, a number of instances of process uncertainty (n=23; 6.0% of all identified uncertainties) were identified, suggesting that the corresponding WOE assessments (in which process uncertainty features) are being performed without an appropriate level of scientific evidence, either on purpose or from a lack of choice.

Whilst only one other correlation of any great strength was identified in this research (between human (variability) uncertainty and expert opinion; ρ=0.55), the existence of so many positive correlations (103 out of the possible 315) between identified uncertainties and the LOEs potentially carries more meaning. The studied evidence base consists of WOE assessments from a vast array of scientific fields, each relying on different LOEs to differing degrees. The presence of numerous weak associations between LOEs and identified uncertainties within the evidence base served to dilute the strength (in a statistical context) of other more meaningful relationships. By conducting similar research on the frequency and statistical dependencies between uncertainties and evidence utilised, but in a more specific research area (where the full range of uncertainties can exist and the LOEs in operation are more directed), the outlined relationship between uncertainties and evidence may increase in worth. Should stronger individual relationships between uncertainties and LOEs be identified, the opportunity might exist to reverse them, utilising the different evidence types (and sub-types) as catalysts to predict the occurrence of different uncertainties. For example, on the basis of the statistical correlations presented, when utilising expert opinion evidence, human (variability) uncertainty is expected to exist. Such guidance could be implemented by extending the traditional typology format (i.e. lists of terminology and associated definitions), and would be a valuable tool, particularly in the formative stages of uncertainty analysis.

4.5 Conclusion

Uncertainty typologies act both as communication tools and aids in the identification process. The categorisations and definitions presented must therefore be comprehensive and reliable. However, existing typologies have been found to be lacking in a number of ways. This chapter presented a typology of uncertainties based, for the first time, on the analysis of a large evidence base, namely peer-reviewed environmental WOE assessments. The new typology (see Table 4.2), which consists of 7 types of uncertainty (data, language, system, extrapolation, variability, model, and decision) and 20 related sub-types, has resolved several key issues surrounding existing typologies, including research domain transferability and content reliability issues. In addition, this chapter has shown that whilst the majority of techniques used to manage these uncertainties were used appropriately, some assessors are needlessly impacting the validity of their results by ignoring uncertainty altogether. Finally, the relationships that exist between uncertainties and several aspects of the assessment process have been explored, highlighting a strong relationship between uncertainty and the evidence utilised in assessments, which offers opportunities moving forward with respect to improving the uncertainty identification process.

Chapter 5: An uncertainty identification system for environmental risk assessments

5.1 Introduction

Typologies can provide reliable characterisations of potential uncertainties for consideration during ERAs (see Chapter 3).However, the differing abilities and experience levels of ERA practitioners can result in these typologies being used inconsistently (Gillund *et al.* 2007; Knol *et al.* 2009a), which can potentially cause uncertainties in ERAs to go unidentified (EEA 2007; Hart *et al.* 2007; Dale *et al.* 2008). Critically, a full framework for comprehensively describing uncertainty – which would offer more specific guidance regarding the identification of uncertainties in ERAs than is found in a typology – is still lacking (Siegel 2010). Building on the novel evidence-based typology (described in Chapter 4), this chapter introduces a system that uses and extends the existing typology format to aid uncertainty identification in ERAs.

5.2 Method

5.2.1 Overview

The method of creating and validating system maps to be used as the basis for expert engagement, predominantly to elicit views about risks and uncertainties, is gaining in popularity within the risk community (Krayer von Krauss *et al.* 2004; Gillund *et al.* 2008; Krayer von Krauss *et al.* 2008; Ravnum *et al.* 2012; Smita *et al.* 2012; Zimmer *et al.* 2012). This approach, together with a combination of literature-based research and expert elicitation was used to develop an uncertainty identification system for environmental risk assessments (UnISERA). As an overview (Figure 5.1), the main methodological process began by creating and validating a template (analogous to a system map) containing the important aspects of a generic ERA, followed by the selection of suitable risk domains on which to base UnISERA (see Section 5.3). Then, for each selected risk domain, an evidence base of relevant ERAs was established, from which a dominant risk relationship was selected as the focus. A domain-specific version of the generic ERA template was then created for each risk relationship and subsequently validated. Elicitation systems were designed, using the information contained within the templates, and executed using domain experts. The results of the elicitations in the three risk domains were then aggregated to create a single representation of the levels, natures, and locations of uncertainties within ERAs. Validation against a separate risk domain followed (see Chapter 6). This method is explored in detail in the following subsections.

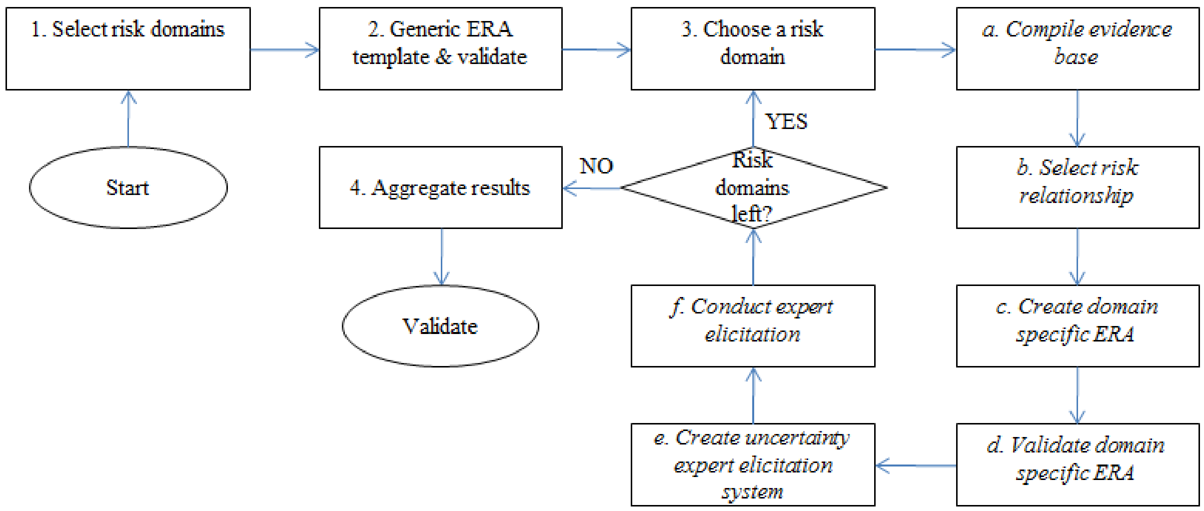


Figure 5.1 Overview of the methodological approach adopted in creating the uncertainty identification system for environmental risk assessments.

5.2.2 Generic ERA template

The objective of this chapter is to describe uncertainty across ERAs in as much detail as possible. Therefore, the structural integrity of the uncertainty-based expert elicitation system requires that all features of conventional ERAs be included. These features are described in several published peer-reviewed and regulatory literature sources (see Section 5.3). Examination of these sources enabled the creation of a template conveying the important features of ERAs, here termed the generic ERA template, version 1. Validation of this initial template was performed in two separate rounds using the views and opinions of experts based in the wide domain of ERA.

Experts were sourced using the online academic search engine Scopus, chosen since it covers a wider journal range (including academic and industrial trade journals) than any other search engine (Falagas *et al.* 2008), using the search term *'risk assessment* (in article title) AND *ecological* OR *environmental* OR *human* (in keywords)'. Results were limited to the last 10 years, in an attempt to identify all experts still active in this research domain. In-built filtering within Scopus was used to remove obviously non-relevant literature, with further filtering, to ensure that the returned sources related to ERAs, performed using the information within titles and abstracts. The remaining records were ordered by decreasing citation count, and the contact details of the first 2,000 records were exported. This threshold was set to allow for a 50% redundancy in records (due to outdated contact details, duplicates, or retired individuals), and a 5-10% response rate from the remainder (Speirs-Bridge *et al.* 2010), ideally resulting in 50-100 responses. The email addresses of the corresponding authors were extracted from each record, and duplicate addresses were removed.

For the first round of validation, the generic ERA template, version 1, was sent to every other contact in the compiled list (i.e. all odd entries), along with explanatory information about the research project and details of the validation request. Experts were asked to validate the phases (e.g. problem formulation), sub-phases (e.g. defining the conceptual model), and tasks (e.g. identifying sources, pathways, and receptors) contained within the template. A deadline of four weeks was given, after which the views of the experts were collated, and alterations to the generic ERA template, version 1, were made for suggestions where two or more experts agreed. Completion of the first stage of validation yielded the generic ERA template, version 2.

The procedure was repeated for the second round of validation, with the generic ERA template, version 2, sent to the remaining experts in the contacts list (i.e. all even entries). Completion of the second stage of validation yielded the generic ERA template, version 3, which was used as the structure for the domain-specific ERA templates.

5.2.3 Domain-specific ERA templates

Domain-specific versions of the validated generic ERA template (version 3) were created for each of the three test case studies. A separate evidence base of peer-reviewed ERAs was compiled, using Scopus, for each of these three case studies, using the search term *'risk assessment* (in keywords) AND *domain* (in keywords)', where *domain* represents the name of the subject matter, and changes depending on the evidence base being built. Results were not restricted temporally (e.g. the last 10 years), but were assessed for relevance using the procedure described in Section 5.2.2.

The ERAs within each of the three case study evidence bases were analysed separately, with all risk relationships (i.e. sources, pathways, and receptors), recorded. A single risk relationship was selected for each case study, on the basis of the most frequently occurring set of S-P-Rs, and the three evidence bases were updated to contain the corresponding ERAs only. Information contained within these ERAs was then used to populate the generic ERA template, version 3, thus creating three domain-specific ERA templates, one representing each risk relationship. The information within the domain-specific ERA templates was then used to create three separate uncertainty-based expert elicitation systems.

5.2.4 Uncertainty-based expert elicitations

The expert elicitation systems were designed in Microsoft Excel 2007 (Microsoft Corporation, Redmond WA) using form controls and macros, and were distributed to experts via email, to be completed remotely and returned before a set deadline. Ethics approval was granted by Cranfield University. The elicitations were conducted according to recommended procedures (Slottje *et al.* 2008; US EPA 2009; Knol *et al.* 2010), following a seven-step elicitation procedure derived from existing methodologies (Knol *et al.* 2010).

Step 1: characterisation of uncertainties

The elicitation used the typology described in Chapter 4, which consisted of three elements within the nature dimension (epistemic, aleatory, and combined) and seven elements across the location dimension (data, language, system, variability, extrapolation, model, and decision). Sub-types of location-based uncertainties were not included, in order to keep the elicitation completion time to a realistic length. The level dimension was depicted as a range of integers (from zero to 10, where zero represents a deterministic understanding of the uncertainty and 10 represents total ignorance to it), consistent with similar research (Krayer von Kraus *et al.* 2004; Figure 5.2; see Section 2.7 for more information).

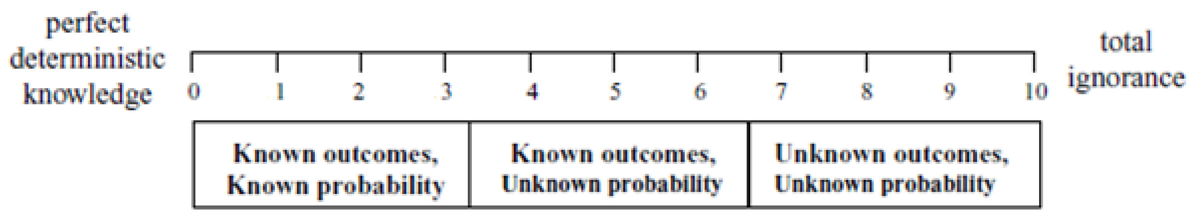


Figure 5.2 The quantitative scale used to assess the level of uncertainty, in the view of the expert (Krayer von Kraus *et al.* 2004).

Step 2: scope and format of the elicitation

The elicitations, in the form of computerised questionnaires implemented in Microsoft Excel 2007, were sent to the experts via email to be completed remotely. Experts were approached as individuals rather than as groups. The elicitation design is discussed further in step 4.

Step 3: selection of experts

The experts selected to participate in this elicitation are the same that were invited to validate the domain-specific ERA templates, unless they specifically stated a wish not to participate, and are therefore considered to be subject-matter experts, drawn from academia, industry, and regulatory agencies. In addition, the credentials of an expert who completed an elicitation were checked to ensure recent and extended involvement in the relevant domain of the elicitation.

Step 4: design of the elicitation protocol

The elicitation is organised according to the phases, tasks, and sub-tasks contained within the validated generic ERA template (see Section 5.2.2), all of which were contextualised for the experts using corresponding information from the relevant domain-specific ERA template (see Section 5.2.3). Experts were asked to assess four aspects for every task within each ERA phase:

1) using tick-box controls, whether the task is necessary in an ERA of the domain, thereby providing an extra validation of the elicitation contents. If deemed unnecessary, experts were able to move to the next task;
2) using slide bars, the level of the uncertainty associated with performing the task (on a scale of zero to 10; Figure 5.2);
3) Using tick-box controls, the nature of the uncertainty associated with performing the task (epistemic, aleatory, or combined); and
4) using tick-box controls, the location(s) of the uncertainty associated with performing the task (data, language, system, variability, extrapolation, model, or decision).

The total number of tasks evaluated by an expert during the elicitation was the same as the total number of tasks contained within the relevant validated domain-specific ERA template.

In order to ensure that experts understood the uncertainty-based concepts, they were asked to complete a practice exercise prior to starting the elicitation, based on the introduction of a DNA vaccine into aquaculture (Gillund *et al.* 2008). Experts were provided with an overview of the topic, some background information, an explanatory figure, and a set of instructions. The practice task, which consisted of five questions, followed the same format as the main elicitation section, helping to familiarise experts with the structure. The answers provided by experts, relating to the level and nature of uncertainty, were compared to the 'control' set from the original elicitation (Gillund *et al.* 2008), in which the location of uncertainty was not assessed. Experts involved in this elicitation were considered to understand and be able to assess the level of uncertainty (i.e. not be overly optimistic or pessimistic when faced with a scenario) if their averaged results were within ±50% of that of the control group. With regard to the nature of the uncertainty, experts were expected to agree with the control group to a minimum level of 60% (i.e. three out of five questions). Provided that these two criteria were met, the judgements within the completed corresponding elicitations were treated as valid. A thorough written or verbal communication was held with experts who failed to complete the practice exercise, to satisfy that they understood the associated uncertainty-based concepts.

All questions (namely the tasks) within an elicitation were worded in a consistent manner, using the validated common terminology within the generic and relevant domain-specific ERA template. Potential bias in responses was reduced by providing the experts with personal and professional anonymity throughout.

Step 5: preparation of the elicitation session

The distribution of domain-specific ERA templates to all potential experts, for the purpose of validation, ensured their proper preparation. Both the validated generic ERA and relevant domain-specific ERA templates were provided for experts to view as part of the introductory information within the elicitation system.

Step 6: elicitation of expert judgements

The elicitation was organised into three main sections: an introduction, which included an elicitation overview and background on uncertainty dimensions and the ERA process (Appendix B); instructions on how to complete the elicitation, including the method used to assess the levels, natures, and locations of uncertainty, as well as the pre-elicitation practice exercise (Appendix C); and the main elicitation, further separated into the four phases of an ERA.

Step 7: possible aggregation and reporting

Due to the stringent selection criteria used here, the responses of all experts were considered to be of equal importance. Therefore, equal weights methods were chosen (Clemen and Winkler 1999; Slottje *et al.* 2008). Specifically, the responses for the levels of uncertainty were aggregated using measures of central tendency, with the natures and locations of uncertainty combined into occurrence percentages (the following sub-section contains more information about the types of data gathered). The individual responses associated with each distinct ERA task were aggregated and included within UnISERA provided that that task featured in at least two of the three elicitation case studies. This ensured that UnISERA was inclusive and representative and not, in any part, restricted to observations from a single subject domain.

5.2.5 Data analysis

Data collected from the expert elicitations were analysed for relationships and trends. All ERA tasks within the elicitations of each case study had two kinds of data associated with them: the level of uncertainty (measured using slide bars) was represented through integer values in the range of zero to 10, whilst the nature and location of uncertainty (measured using tick box controls) were treated as binary values. The data from completed elicitations were extracted and stored in a separate spreadsheet, alongside an assigned expert ID.

With respect to the level of uncertainty associated with each ERA task, both within individual case studies and for the UnISERA (i.e. combined) data, relationships were evaluated using central tendencies, variations from the central tendency, and the high-low ranges of responses. With respect to the level of uncertainty associated with the ERA tasks in UnISERA (i.e. across the two or three case studies in which the tasks feature), relationships were evaluated statistically through appropriate comparison of the measures of central tendency. The data for the natures and locations of uncertainty, for both the individual case studies and UnISERA, were converted from binary values to occurrence percentages.

Selection of appropriate measures or statistical tests to analyse the central tendencies of the levels of uncertainty were dictated by the distributions of the datasets. Here, suitable parametric (mean and standard deviation, two-sample t-test or ANOVA) or non-parametric (median and inter-quartile range, Mann-Whitney or Kruskal-Wallis) measures and statistical tests were applied to the datasets, as highlighted in the relevant results section (Sections 5.4-5.7), based on an assessment of their normality. All statistical tests were performed using SPSS v19 (SPSS Inc., Chicago IL).

Categorical (i.e. binary) data does not need to be assessed for normality, and was considered here to be non-parametric due to the small size of the datasets analysed. Furthermore, binary data does not have a measure of central tendency and cannot be evaluated in a statistically similar way to the integer values in the level dimension.

5.3 Results 1: risk domain selection and generic ERA template

5.3.1 Risk domain selection

There are many potential risk domains upon which UnISERA could be based, and as such the selection of three domains contained some subjectivity. However, strict criteria were also adhered to. In the context of this research, a suitable risk domain is one which:

- has a large amount of associated empirical evidence;

- ERAs are prevalent throughout;
- is relevant to UK-based practitioners (i.e. falls within Defra's landscape of risk concerns); and
- environmental uncertainty is present.

The three risk domains of focus, based on these criteria, were selected as genetically modified higher plants, particulate matter, and pesticides (see Section 2.3.3 for background information).

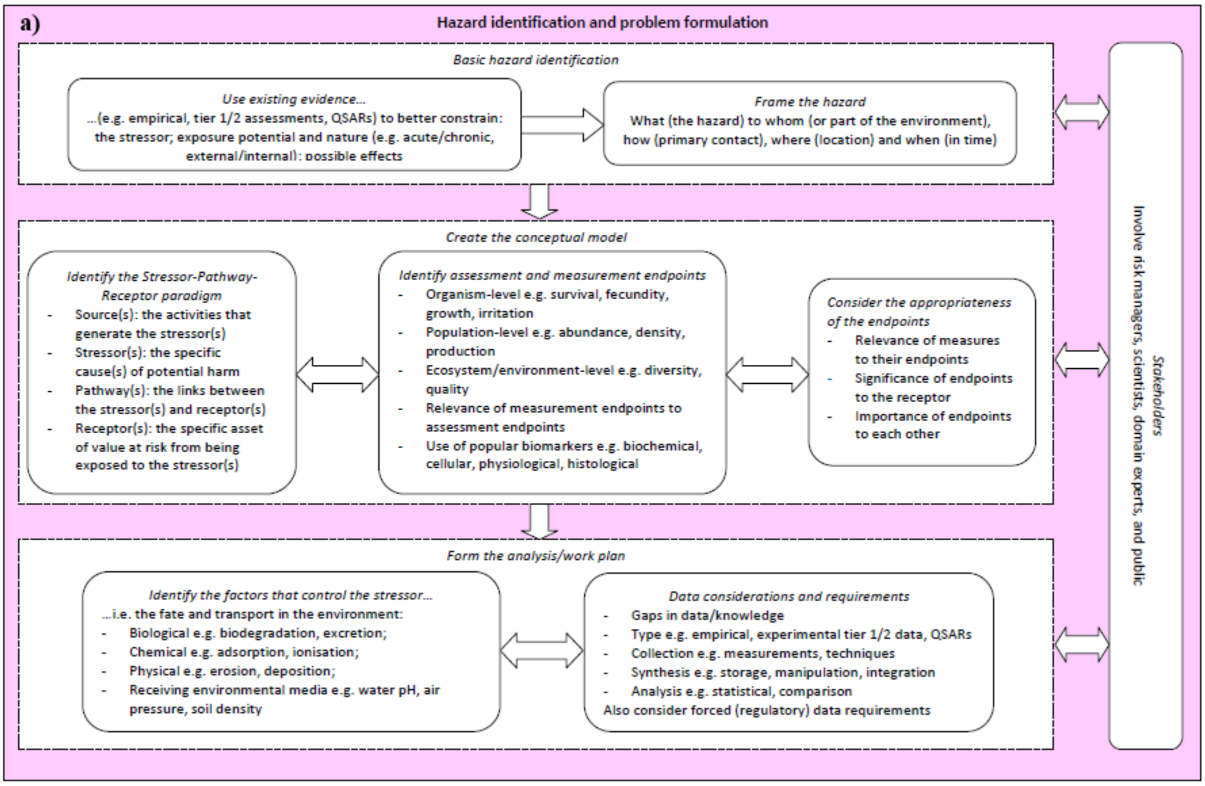
5.3.2 Generic ERA template creation and validation

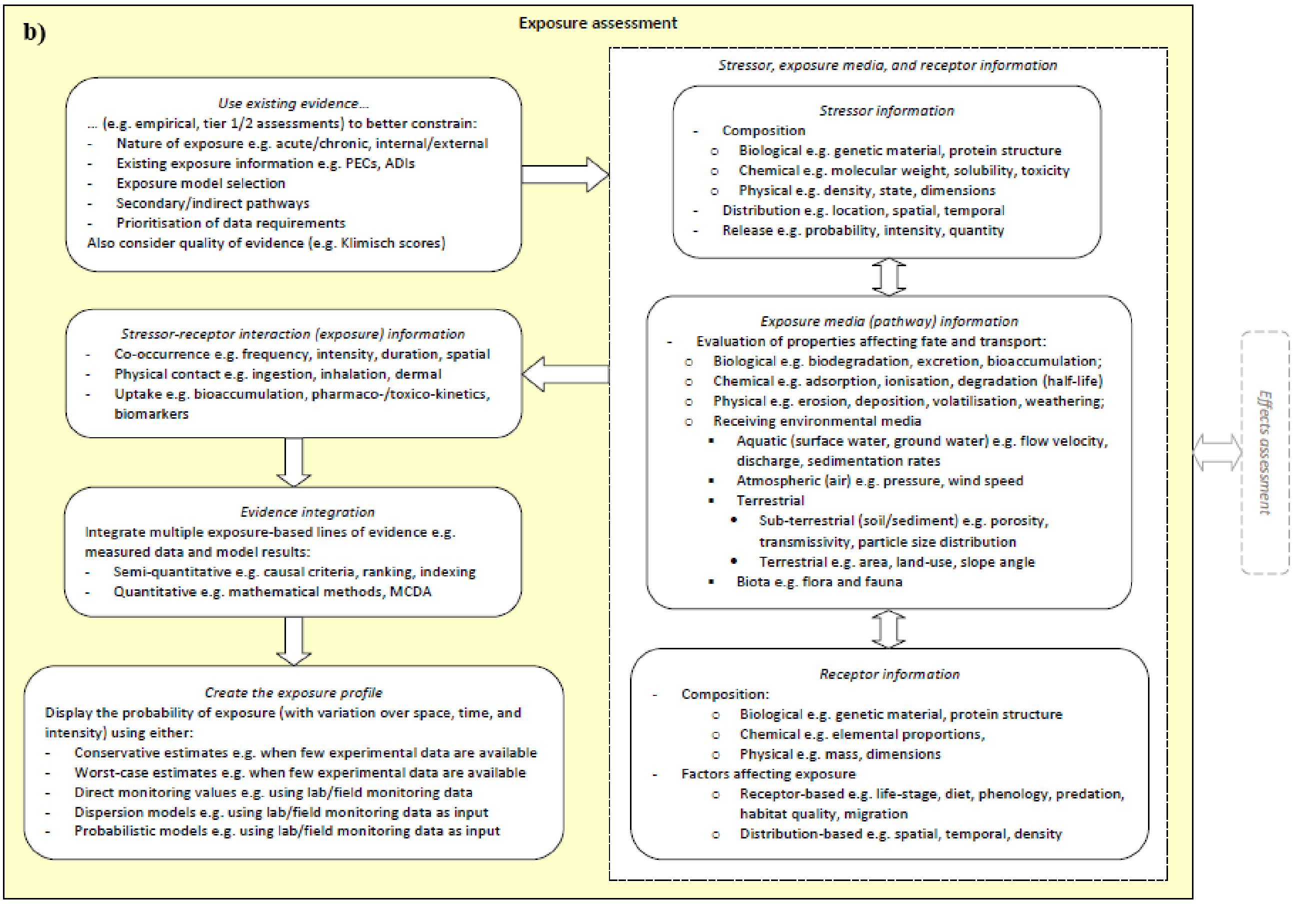
The generic ERA template, version 1 (Appendix D) was created using the information within several academic and grey literature sources (US EPA 1992; Suter 1996; US EPA 1998; Fairman *et al.* 1998; DETR/EA and IEH 2000; DHA 2002; US EPA 2003; Landis 2005; Beer 2006; Briggs 2008; Defra 2011). The 35 experts (21 from academia, 4 from industry, 10 from regulatory agencies) involved in the first round of validation provided comments regarding the correctness and completeness of this template (Table 5.1; Supplementary Material A).

Table 5.1 The number of comments received from experts involved in the two rounds of validation for the generic ERA template (versions 1 and 2), and the number of changes made to the templates (with example changes), organised by ERA phase. The full list of comments is provided in Supplementary Materials A and B.

ERA phase	# comments	# changes	Example change
		Validation round 1	
Problem formulation	28	8	• Removed the 'risk screening' box • Differentiated the 'source' and 'stressor' terms
Exposure assessment	22	4	• Included 'characteristics of the receptor' as a data collection consideration • Included an 'evidence integration' box
Effects assessment	35	7	• Modified the 'create the stressor-response profile' to include variation over time, space and intensity • Removed 'consider the relevance of endpoints' in stressor-receptor relationship analysis step
Risk characterisation	26	4	• Included options to aggregate risk levels • Included a 'risk evaluation' box, including assessing the significance of the risk
		Validation round 2	
Problem formulation	11	1	• Expanded the 'relevance of endpoints' section to include significance and importance measures
Exposure assessment	10	1	• Differentiated between dispersion and probabilistic models when creating the exposure profile(s)
Effects assessment	11	0	No changes made
Risk characterisation	5	1	• Included regulatory, stakeholder, and experimentally-derived thresholds when assessing the significance of the risk

The implementation of these comments enabled the creation of the generic ERA template, version 2 (Appendix E). Validation of this template, through comments (Table 5.1; Supplementary Material B) provided by 13 different experts (9 from academia, 2 from industry, 2 from regulatory agencies), enabled the creation of the generic ERA template, version 3 (Figure 5.3).


a)
Hazard identification and problem formulation
Basic hazard identification
Use existing evidence…
…(e.g. empirical, tier 1/2 assessments, QSARs) to better constrain: the stressor; exposure potential and nature (e.g. acute/chronic, external/internal); possible effects
Frame the hazard
What (the hazard) to whom (or part of the environment), how (primary contact), where (location) and when (in time)
Create the conceptual model
Identify the Stressor-Pathway-Receptor paradigm
- Source(s): the activities that generate the stressor(s)
- Stressor(s): the specific cause(s) of potential harm
- Pathway(s): the links between the stressor(s) and receptor(s)
- Receptor(s): the specific asset of value at risk from being exposed to the stressor(s)
Identify assessment and measurement endpoints
- Organism-level e.g. survival, fecundity, growth, irritation
- Population-level e.g. abundance, density, production
- Ecosystem/environment-level e.g. diversity, quality
- Relevance of measurement endpoints to assessment endpoints
- Use of popular biomarkers e.g. biochemical, cellular, physiological, histological
Consider the appropriateness of the endpoints
- Relevance of measures to their endpoints
- Significance of endpoints to the receptor
- Importance of endpoints to each other
Form the analysis/work plan
Identify the factors that control the stressor…
…i.e. the fate and transport in the environment:
- Biological e.g. biodegradation, excretion;
- Chemical e.g. adsorption, ionisation;
- Physical e.g. erosion, deposition;
- Receiving environmental media e.g. water pH, air pressure, soil density
Data considerations and requirements
- Gaps in data/knowledge
- Type e.g. empirical, experimental tier 1/2 data, QSARs
- Collection e.g. measurements, techniques
- Synthesis e.g. storage, manipulation, integration
- Analysis e.g. statistical, comparison
Also consider forced (regulatory) data requirements
Stakeholders
Involve risk managers, scientists, domain experts, and public


b)
Exposure assessment
Use existing evidence…
… (e.g. empirical, tier 1/2 assessments) to better constrain:
- Nature of exposure e.g. acute/chronic, internal/external
- Existing exposure information e.g. PECs, ADIs
- Exposure model selection
- Secondary/indirect pathways
- Prioritisation of data requirements
Also consider quality of evidence (e.g. Klimisch scores)
Stressor-receptor interaction (exposure) information
- Co-occurrence e.g. frequency, intensity, duration, spatial
- Physical contact e.g. ingestion, inhalation, dermal
- Uptake e.g. bioaccumulation, pharmaco-/toxico-kinetics, biomarkers
Evidence integration
Integrate multiple exposure-based lines of evidence e.g. measured data and model results:
- Semi-quantitative e.g. causal criteria, ranking, indexing
- Quantitative e.g. mathematical methods, MCDA
Create the exposure profile
Display the probability of exposure (with variation over space, time, and intensity) using either:
- Conservative estimates e.g. when few experimental data are available
- Worst-case estimates e.g. when few experimental data are available
- Direct monitoring values e.g. using lab/field monitoring data
- Dispersion models e.g. using lab/field monitoring data as input
- Probabilistic models e.g. using lab/field monitoring data as input
Stressor, exposure media, and receptor information
Stressor information
- Composition
o Biological e.g. genetic material, protein structure
o Chemical e.g. molecular weight, solubility, toxicity
o Physical e.g. density, state, dimensions
- Distribution e.g. location, spatial, temporal
- Release e.g. probability, intensity, quantity
Exposure media (pathway) information
- Evaluation of properties affecting fate and transport:
o Biological e.g. biodegradation, excretion, bioaccumulation;
o Chemical e.g. adsorption, ionisation, degradation (half-life)
o Physical e.g. erosion, deposition, volatilisation, weathering;
o Receiving environmental media
▪ Aquatic (surface water, ground water) e.g. flow velocity, discharge, sedimentation rates
▪ Atmospheric (air) e.g. pressure, wind speed
▪ Terrestrial
• Sub-terrestrial (soil/sediment) e.g. porosity, transmissivity, particle size distribution
• Terrestrial e.g. area, land-use, slope angle
▪ Biota e.g. flora and fauna
Receptor information
- Composition:
o Biological e.g. genetic material, protein structure
o Chemical e.g. elemental proportions,
o Physical e.g. mass, dimensions
- Factors affecting exposure
o Receptor-based e.g. life-stage, diet, phenology, predation, habitat quality, migration
o Distribution-based e.g. spatial, temporal, density
Effects assessment

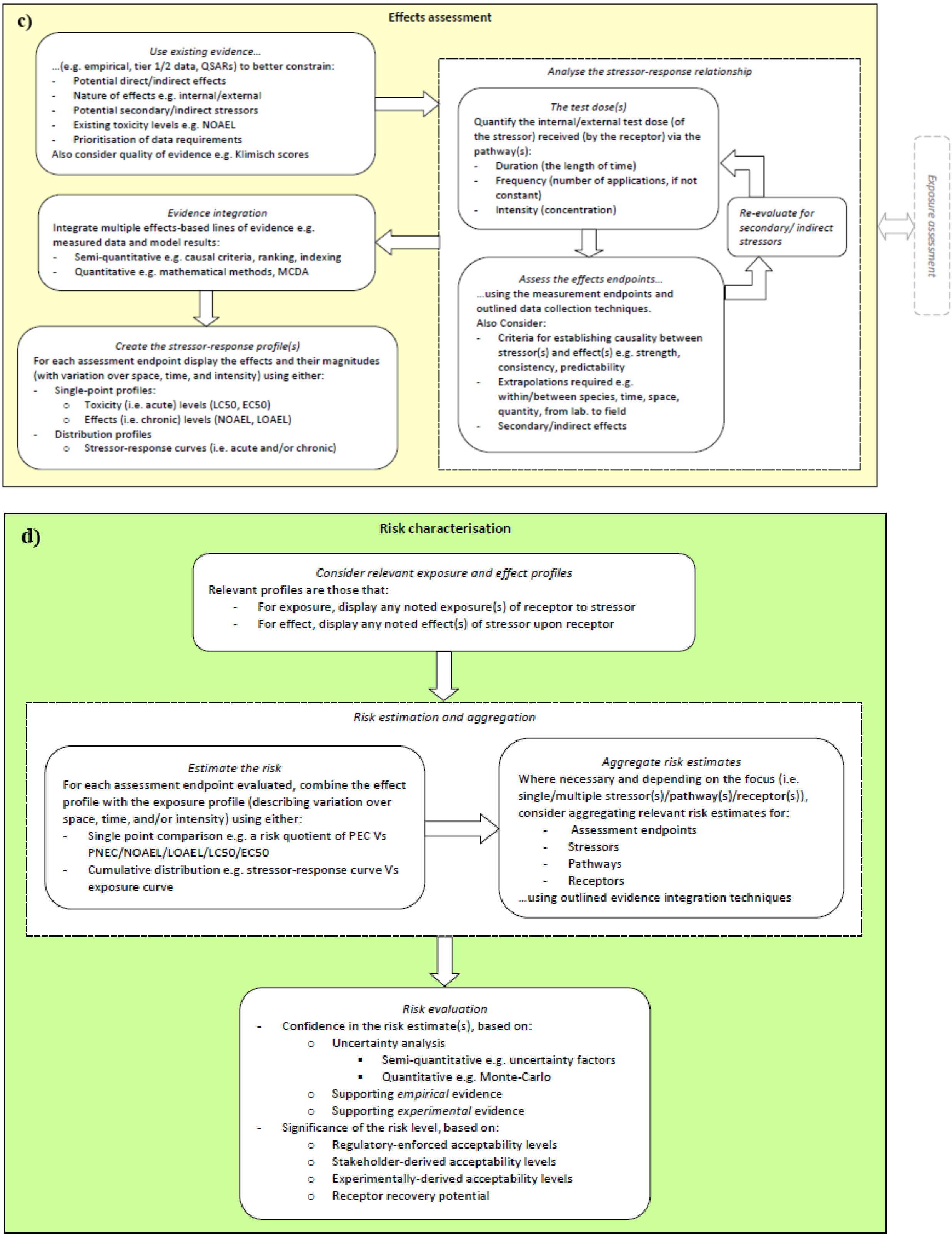


Figure 5.3 The generic ERA template, version 3, created through the expert validation of versions 1 and 2, describing the important aspects within the phases of: a) hazard identification and problem formulation; b) exposure assessment; c) effects assessment; and d) risk characterisation.

The information contained within this generic template, which formed the basis of the domain-specific templates and therefore the expert elicitations, was organised according to the different phases, sub-phases, task groups, and task numbers, and is shown in Table 5.2.

Table 5.2 The 105 ERA tasks within the generic ERA template, version 3, organised by task group, sub-phase, and phase, to be potentially included in the expert elicitation exercises across the three case studies.

ERA Phase	ERA sub-phase	ERA task group	ERA Task number
Problem formulation	Preliminary hazard identification	1. Use available evidence to better constrain…	1-4
		2. Framing the hazard	5-9
	Define the conceptual model	3. Identify the S-P-R paradigm, including…	10-13
		4. Choose assessment and measurement endpoints	14-21
		5. Consider the appropriateness of the endpoints	22-24
	Form the analysis/work plan	6. Identify the factors controlling fate and transport of the stressor	25-28
		7. Identify data considerations	29-32
Exposure assessment	Use available evidence to better constrain…	8. (Use available evidence to better constrain…)	33-37
	Stressor, exposure media, and receptor information	9. Collect information about the stressor's composition	38-40
		10. Collect information about the stressor's distribution	41-42
		11. Collect information about the stressor's release	43-45
		12. Collect information about properties affecting fate and transport	46-53
		13. Collect information about the receptor	54-57
	Evaluate stressor-receptor contact	14. Evaluate co-occurrence for…	58-60
		15. Evaluate…	61-62
	Integrate multiple LOEs using…	16. (Integrate multiple LOEs using…)	63-64

	Create the exposure profile(s) using…	17. (Create the exposure profile(s) using…)	65-69
Effects assessment	Use available evidence to better constrain…	18. (Use available evidence to better constrain…)	70-74
	Analyse the stressor-response relationship	19. Determine the test dose for the…	75-77
		20. Assess effect endpoints	78-85
	Integrate multiple LOEs using…	21. (Integrate multiple LOEs using…)	86-87
	Create stressor-response profile using…	22. Single point or distribution methods showing…	88-91
Risk characterisation	Select relevant profiles…	23. (Select relevant profiles…)	92-93
	Estimate and aggregate risk	24. Estimate risk using…	94-95
		25. Aggregate risk estimates for…	96-99
	Evaluate risk levels	26. Assess confidence in the risk levels using…	100-101
		27. Assess the significance of the risk levels using…	102-105

5.4 Results 2: Case Study 1 (genetically modified higher plants)

5.4.1 Risk relationship selection

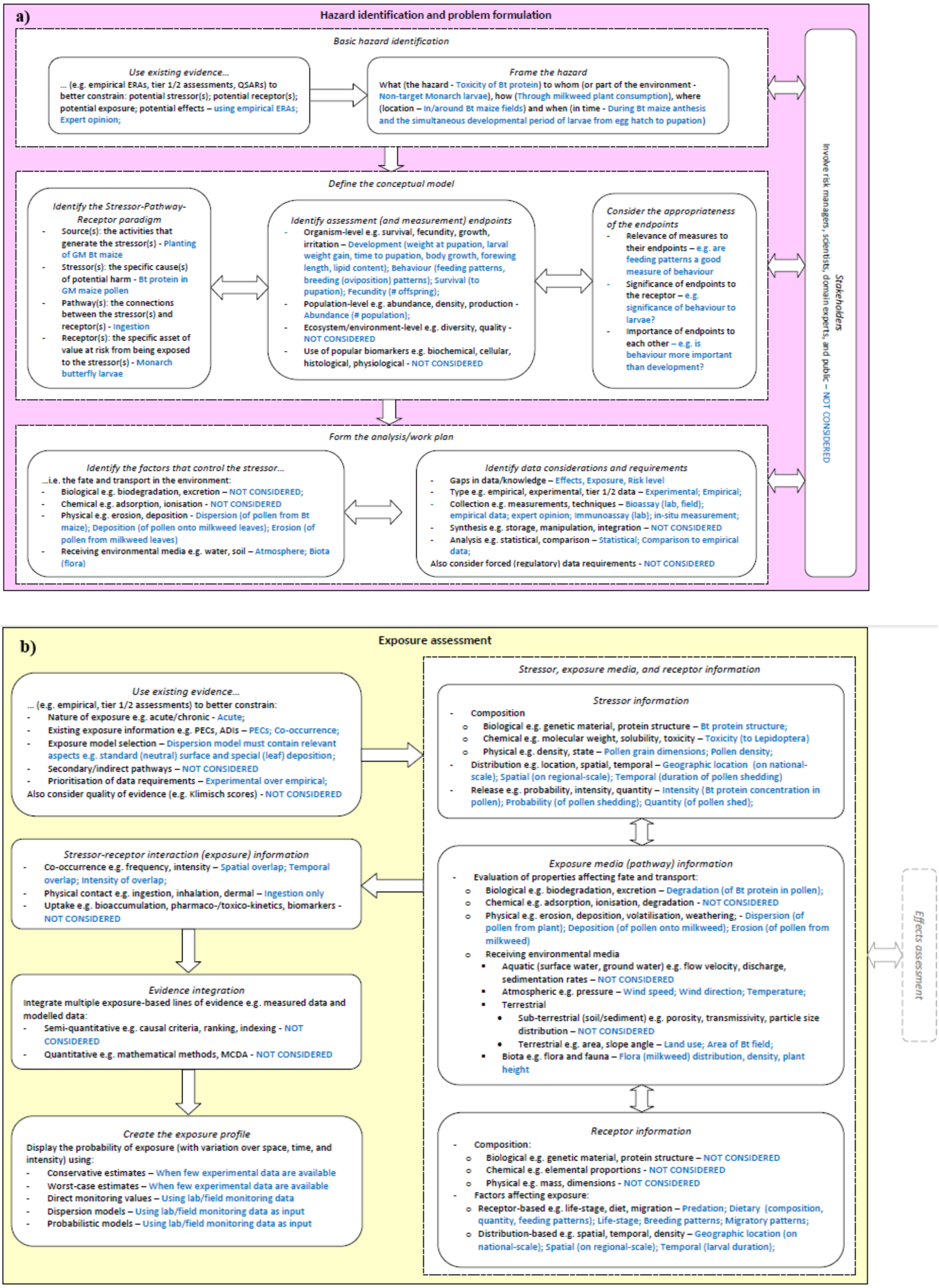
The literature searches, after in-built filtering and relevance-checking, returned 155 peer-reviewed articles, which were further reduced, on the basis of missing information within the articles, yielding a GMHP evidence base of 118 articles. The GMHP evidence base was analysed for its risk relationships, the most frequent of which (n=19) was potential *Bacillus thuringiensis* (Bt) modified maize (*Zea mays*) risk to non-target Lepidoptera. Of these 19 articles, 13 (Losey *et al.* 1999; Jesse and Obrycki 2000; Hellmich *et al.* 2001; Oberhauser *et al.* 2001; Sears *et al.* 2001; Stanley-Horn *et al.* 2001; Zangerl *et al.* 2001; Wolt *et al.* 2003; Dively *et al.* 2004; Anderson *et al.*, 2005; Mattila *et al.* 2005; Gathmann *et al.* 2006; Perry *et al.* 2010) focused on the larvae of Monarch butterflies (*Danaus plexippus* L.) as the receptor of interest. On this basis, 'potential *Bacillus thuringiensis* (Bt) modified maize (*Zea mays*) risk to non-target Monarch butterfly larvae' was selected as the risk relationship for this case study.

5.4.2 ERA template creation and validation

The generic ERA template, version 3, was populated with relevant information from the 13 peer-reviewed articles, forming the Bt-maize risk to Monarch larvae ERA template, version 1 (Appendix F). Validation of this template, through comments (Table 5.3; Supplementary Material C) provided by 7 of the experts in the GMHP evidence base enabled the creation of the Bt-maize risk to Monarch larvae ERA template, version 2 (Figure 5.4). The experts were based in the sectors of academia (n=3), industry (n=1), and regulation (n=3), and reside in Canada (n=1), France (n=1), Germany (n=3), Switzerland (n=1), and USA (n=1).

Table 5.3 The number of comments received from experts involved in the validation of the Bt-maize risk to Monarch larvae ERA template, version 1, and the number of changes made to the template (with example changes), organised by ERA phase. The full list of comments is provided in Supplementary Material C.

ERA phase	# comments	# changes	Example change
Problem formulation	8	2	• Altered task 9 in group 2 to include 'from egg hatch to pupation' • Included expert opinion as a data source in task 31, group 7
Exposure assessment	11	3	• Exposure model requirements (task 35 in group 8) made more specific • Included biological degradation as a factor in task 38, group 9
Effects assessment	5	2	• Included LC/EC toxicity measures in task 73 of group 18 • Removed eclosion effect endpoints from group 20 (and also from group 4 in problem formulation)
Risk characterisation	3	0	No changes made


a)
Hazard identification and problem formulation
Basic hazard identification
Use existing evidence...
... (e.g. empirical ERAs, tier 1/2 assessments, QSARs) to better constrain: potential stressor(s); potential receptor(s); potential exposure; potential effects – using empirical ERAs; Expert opinion;
Frame the hazard
What (the hazard - Toxicity of Bt protein) to whom (or part of the environment - Non-target Monarch larvae), how (Through milkweed plant consumption), where (location – In/around Bt maize fields) and when (in time - During Bt maize anthesis and the simultaneous developmental period of larvae from egg hatch to pupation)
Define the conceptual model
Identify the Stressor-Pathway-Receptor paradigm
- Source(s): the activities that generate the stressor(s) - Planting of GM Bt maize
- Stressor(s): the specific cause(s) of potential harm - Bt protein in GM maize pollen
- Pathway(s): the connections between the stressor(s) and receptor(s) - Ingestion
- Receptor(s): the specific asset of value at risk from being exposed to the stressor(s) - Monarch butterfly larvae
Identify assessment (and measurement) endpoints
- Organism-level e.g. survival, fecundity, growth, irritation – Development (weight at pupation, larval weight gain, time to pupation, body growth, forewing length, lipid content); Behaviour (feeding patterns, breeding (oviposition) patterns); Survival (to pupation); Fecundity (# offspring);
- Population-level e.g. abundance, density, production - Abundance (# population);
- Ecosystem/environment-level e.g. diversity, quality - NOT CONSIDERED
- Use of popular biomarkers e.g. biochemical, cellular, histological, physiological - NOT CONSIDERED
Consider the appropriateness of the endpoints
- Relevance of measures to their endpoints – e.g. are feeding patterns a good measure of behaviour
- Significance of endpoints to the receptor – e.g. significance of behaviour to larvae?
- Importance of endpoints to each other – e.g. is behaviour more important than development?
Form the analysis/work plan
Identify the factors that control the stressor...
...i.e. the fate and transport in the environment:
- Biological e.g. biodegradation, excretion – NOT CONSIDERED;
- Chemical e.g. adsorption, ionisation - NOT CONSIDERED
- Physical e.g. erosion, deposition - Dispersion (of pollen from Bt maize); Deposition (of pollen onto milkweed leaves); Erosion (of pollen from milkweed leaves)
- Receiving environmental media e.g. water, soil – Atmosphere; Biota (flora)
Identify data considerations and requirements
- Gaps in data/knowledge – Effects, Exposure, Risk level
- Type e.g. empirical, experimental, tier 1/2 data – Experimental; Empirical;
- Collection e.g. measurements, techniques – Bioassay (lab, field); empirical data; expert opinion; Immunoassay (lab); in-situ measurement;
- Synthesis e.g. storage, manipulation, integration – NOT CONSIDERED
- Analysis e.g. statistical, comparison – Statistical; Comparison to empirical data;
Also consider forced (regulatory) data requirements - NOT CONSIDERED
Stakeholders
Involve risk managers, scientists, domain experts, and public – NOT CONSIDERED
b)
Exposure assessment
Use existing evidence...
... (e.g. empirical, tier 1/2 assessments) to better constrain:
- Nature of exposure e.g. acute/chronic - Acute;
- Existing exposure information e.g. PECs, ADIs – PECs; Co-occurrence;
- Exposure model selection – Dispersion model must contain relevant aspects e.g. standard (neutral) surface and special (leaf) deposition;
- Secondary/indirect pathways – NOT CONSIDERED
- Prioritisation of data requirements – Experimental over empirical;
Also consider quality of evidence (e.g. Klimisch scores) - NOT CONSIDERED
Stressor-receptor interaction (exposure) information
- Co-occurrence e.g. frequency, intensity – Spatial overlap; Temporal overlap; Intensity of overlap;
- Physical contact e.g. ingestion, inhalation, dermal – Ingestion only
- Uptake e.g. bioaccumulation, pharmaco-/toxico-kinetics, biomarkers - NOT CONSIDERED
Evidence integration
Integrate multiple exposure-based lines of evidence e.g. measured data and modelled data:
- Semi-quantitative e.g. causal criteria, ranking, indexing - NOT CONSIDERED
- Quantitative e.g. mathematical methods, MCDA - NOT CONSIDERED
Create the exposure profile
Display the probability of exposure (with variation over space, time, and intensity) using:
- Conservative estimates – When few experimental data are available
- Worst-case estimates – When few experimental data are available
- Direct monitoring values – Using lab/field monitoring data
- Dispersion models – Using lab/field monitoring data as input
- Probabilistic models – Using lab/field monitoring data as input
Stressor, exposure media, and receptor information
Stressor information
- Composition
o Biological e.g. genetic material, protein structure – Bt protein structure;
o Chemical e.g. molecular weight, solubility, toxicity – Toxicity (to Lepidoptera)
o Physical e.g. density, state – Pollen grain dimensions; Pollen density;
- Distribution e.g. location, spatial, temporal – Geographic location (on national-scale); Spatial (on regional-scale); Temporal (duration of pollen shedding)
- Release e.g. probability, intensity, quantity – Intensity (Bt protein concentration in pollen); Probability (of pollen shedding); Quantity (of pollen shed);
Exposure media (pathway) information
- Evaluation of properties affecting fate and transport:
o Biological e.g. biodegradation, excretion – Degradation (of Bt protein in pollen);
o Chemical e.g. adsorption, ionisation, degradation - NOT CONSIDERED
o Physical e.g. erosion, deposition, volatilisation, weathering; - Dispersion (of pollen from plant); Deposition (of pollen onto milkweed); Erosion (of pollen from milkweed)
o Receiving environmental media
▪ Aquatic (surface water, ground water) e.g. flow velocity, discharge, sedimentation rates – NOT CONSIDERED
▪ Atmospheric e.g. pressure – Wind speed; Wind direction; Temperature;
▪ Terrestrial
• Sub-terrestrial (soil/sediment) e.g. porosity, transmissivity, particle size distribution – NOT CONSIDERED
• Terrestrial e.g. area, slope angle – Land use; Area of Bt field;
▪ Biota e.g. flora and fauna – Flora (milkweed) distribution, density, plant height
Receptor information
- Composition:
o Biological e.g. genetic material, protein structure – NOT CONSIDERED
o Chemical e.g. elemental proportions - NOT CONSIDERED
o Physical e.g. mass, dimensions - NOT CONSIDERED
- Factors affecting exposure:
o Receptor-based e.g. life-stage, diet, migration – Predation; Dietary (composition, quantity, feeding patterns); Life-stage; Breeding patterns; Migratory patterns;
o Distribution-based e.g. spatial, temporal, density – Geographic location (on national-scale); Spatial (on regional-scale); Temporal (larval duration);
Effects assessment

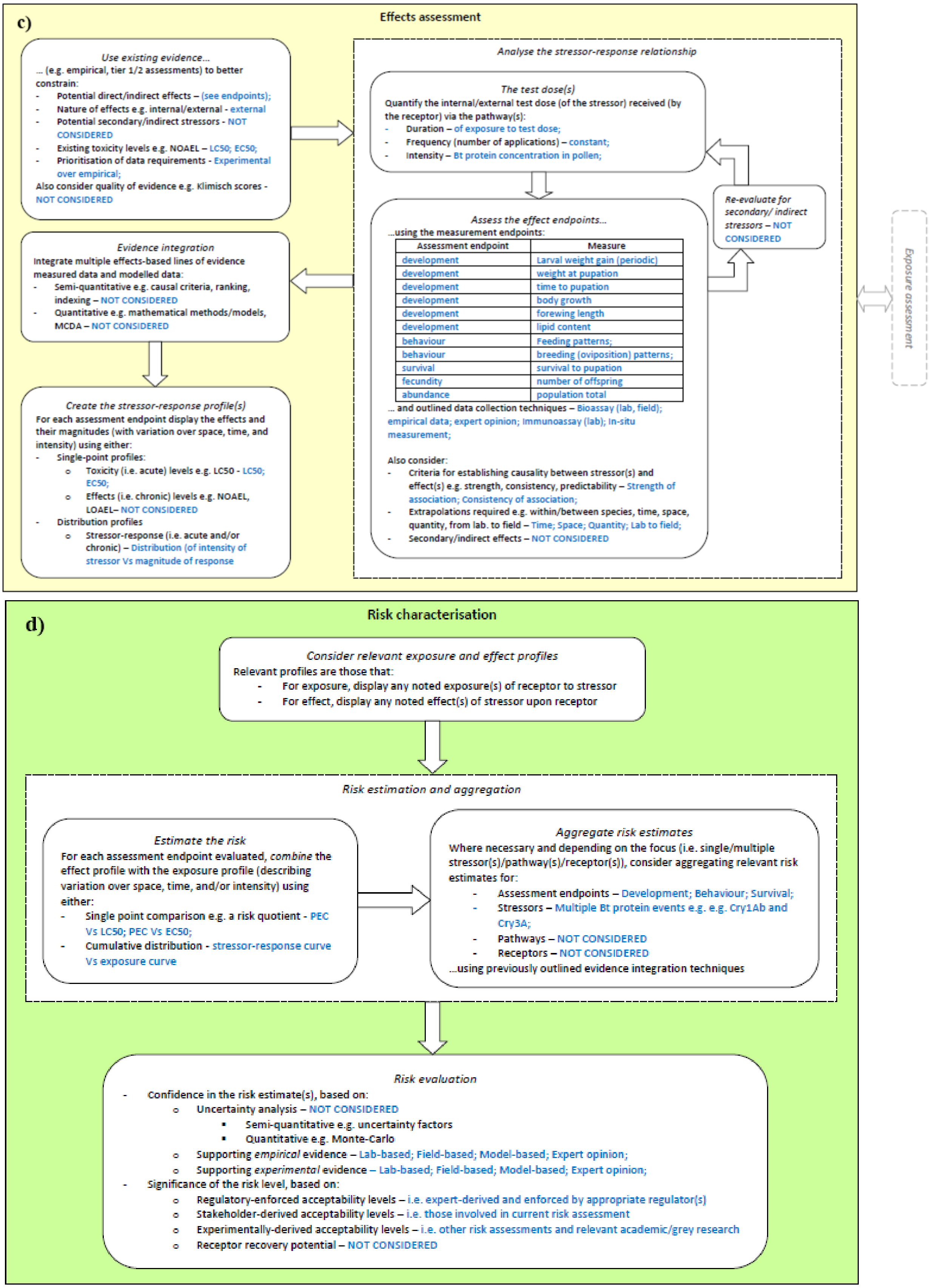


Figure 5.4 The Bt-maize risk to Monarch larvae ERA template, version 2, created through the expert validation of version 1, describing the important aspects within the phases of: a) hazard identification and problem formulation; b) exposure assessment; c) effects assessment; and d) risk characterisation.

5.4.3 *Expert elicitation exercise*

Five experts participated in the Bt-maize risk to non-target Monarch butterfly larvae elicitation exercise (hereafter referred to as Case Study 1; Table 5.4), with each expert assessing 82 separate ERA-based tasks (27 in problem formulation, 28 in exposure assessment, 16 in effects assessment, and 11 in risk characterisation) for the levels, natures, and locations of associated uncertainty.

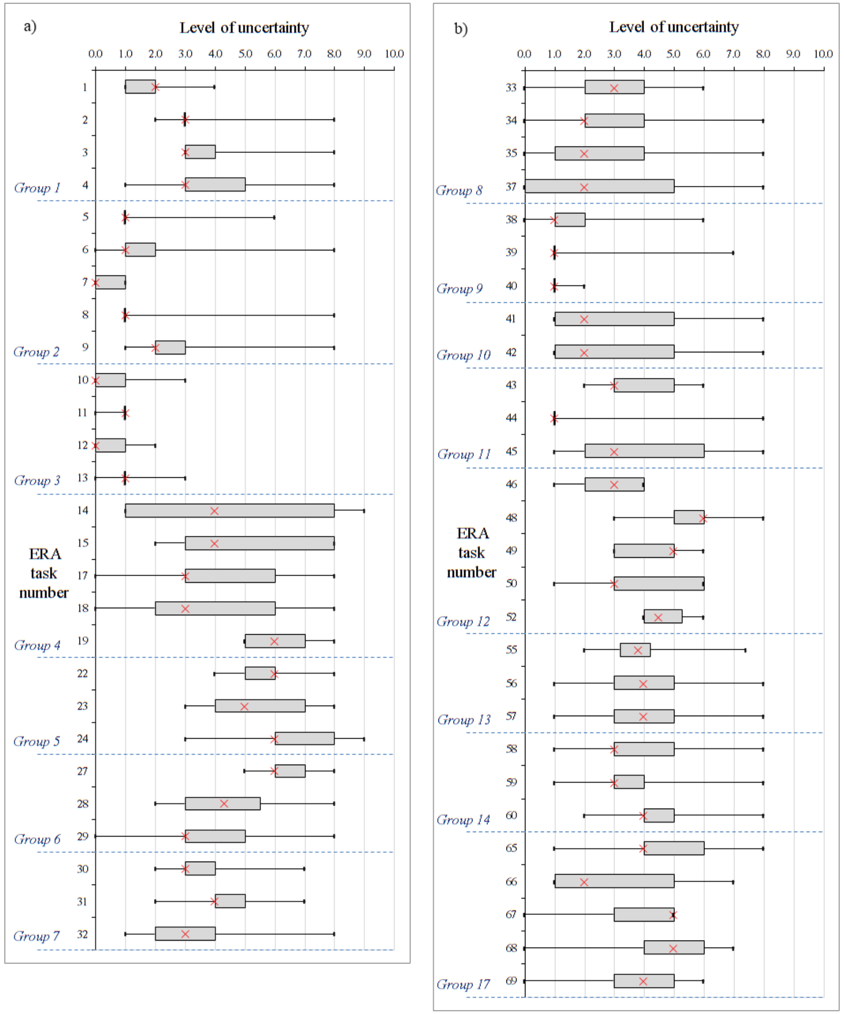
Table 5.4 Professional sectors and countries of residence of the experts (n=5) involved in the uncertainty-based elicitation exercise for Case Study 1. Results from the practice exercise are also included, which show the agreement between the experts and the control group with regard to the level (% above or below the control group mean) and nature (% agreement) of uncertainty communicated.

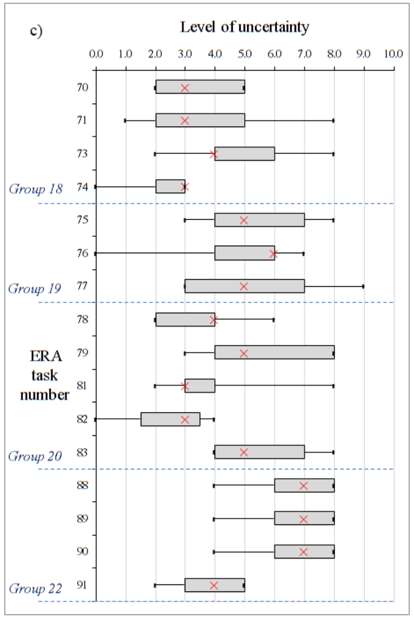
Expert ID	Sector	Country of residence	Level	Nature
1	Academia	South Africa	+19.2%	80%
2	Regulatory	Germany	-	-
3	Academia	United States	-	-
4	Regulatory	United Kingdom	+7.7%	60%
5	Industry	Germany	-4.0%	60%

The 23 tasks not included in this case study (from the 105 in the generic ERA template, version 3) are shown in Section 5.7. Data relating to the level dimension were treated as non-Gaussian after assessment of the mean, median, and mode values; central tendency and spread were measured using median values and inter-quartile ranges (IQRs), respectively.

Problem formulation

Across all aspects, problem formulation had a median level of uncertainty of 3.0 (Figure 5.5a).





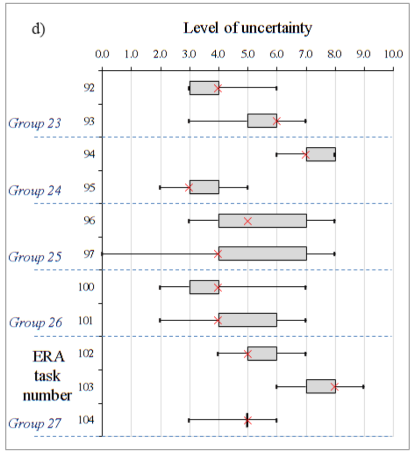


Figure 5.5 The level of uncertainty communicated by the experts (n=5) involved in Case Study 1, across the 82 assessed ERA tasks, organised into the ERA phases of a) problem formulation, b) exposure assessment, c) effects assessment, and d) risk characterisation, and described using median values (red crosses), inter-quartile ranges (boxes), and low-high range values (dashed lines) on a 0 (representing determinism) to 10 (representing total ignorance) scale. The ERA tasks are separated into the groups listed in Table 5.2.

The group of tasks with the highest median level (6.0; group 5) asked experts to consider the appropriateness of assessment and measurement endpoints, whilst the group with the lowest level (0.5; group 3) involved identifying the S-P-R paradigm. Two of the tasks within group 5 also shared the highest levels of uncertainty seen across the individual tasks in this phase (6.0; tasks 22 and 24). Conversely, there were three tasks for which experts communicated median level of determinism (0.0; tasks 7, 10 and 12), two of which belonged to group 3. Groups 2 and 3 saw the highest degree of expert agreement, with IQRs of 0.0 communicated in four of the nine tasks (numbers 5, 8, 11 and 13) in these two groups. Overall expert agreement in this phase was also high, with a median IQR of just 1.0 across all tasks.

The dominant nature of uncertainty was the combined epistemic and aleatory category with a median value of 60%, with either of the other two options preferred just three times (tasks 6, 13 and 24; Table 5.5).

Table 5.5 Median occurrence rates (%) for the individual natures and locations of uncertainty provided by experts (n=5) in Case Study 1, organised by ERA phase and showing the highest rates(s; of at least 50%) for the nature (shaded blue) and location (shaded red) of uncertainty associated with each ERA phase (modal values are included for comparison; median occurrence rates on a task-by-task basis are shown in Appendix G).

ERA phase	Nature of uncertainty (%)			Location of uncertainty (%)						
	Epistemic	Aleatory	Combined	Data	Language	System	Variability	Extrapolation	Model	Decision
Problem formulation median	20	0	60	60	20	20	20	20	40	40
Problem formulation mode	20	0	60	60	0	20	60	20	40	20
Exposure assessment median	20	20	60	60	0	20	60	20	30	0
Exposure assessment mode	20	0	60	60	0	20	60	20	20	0
Effects assessment median	20	20	60	60	0	10	60	30	30	0
Effects assessment mode	20	20	60	60	0	0	60	0	20	0
Risk characterisation median	20	0	80	40	0	40	60	60	40	60
Risk characterisation mode	20	0	80	40	0	20	60	60	40	60
Overall median	20	20	60	60	0	20	60	20	40	20
Overall mode	20	0	60	60	0	20	60	20	20	0

Only one location-based uncertainty had a median occurrence of above 50%, that of the data category (median 60%; Table 5.5). However, unlike in the other phases, all seven categories had a median (but not modal) occurrence of at least 20%.

Exposure assessment

Exposure assessment had a median level of uncertainty of 3.0 (Figure 5.5b), the same as the previous phase of problem formulation. However, the group with the highest median level, which concerned the collection of fate and transport information, was lower than in problem formulation (4.5; group 12), and the group with the lowest median level, which involved collecting information about the stressor's composition, was a little higher (1.0; group 9). Tasks 88, 89 and 90 had the highest individual levels (6.0), and tasks 38, 39, 40 and 44 the lowest (1.0). Experts generally showed a high level of agreement, particularly for tasks 39, 40 and 44 (IQR of 0.0), with an overall IQR of 2.0.

Whilst the dominant nature of uncertainty was again the combined category, with a median occurrence of 60%, experts deemed that the uncertainty associated with eight of the 32 assessed tasks were of a different nature (Table 5.5). Specifically, three of the four tasks in group 8, which concerns using existing evidence at the beginning of the phase to better constrain certain aspects, were epistemic in nature (median of 60%), whilst all three of the tasks in group 11, which relates to collecting information about the stressor's release, were aleatory (60% median).

Data and variability were the most frequent uncertainty locations in exposure assessment (median 60%; Table 5.5). Group 11, which was largely aleatory in nature, showed a particularly high proportion of variability (median 80%), but group 8, which was largely epistemic, highlighted an equal presence of data and system uncertainty (median 60%).

Effects assessment

Effects assessment had a median level of uncertainty of 4.5 (Figure 5.5c), higher than the previous two phases. The group of tasks that concerned creating stressor-response profiles contained the highest level (7.0; group 22), whilst using existing information to better constrain certain aspects at the beginning of the phase had the lowest associated level (3.0; group 18). These two groups also contained the individual tasks with the highest (7.0; tasks 88, 89 and 90) and lowest (3.0, tasks 70, 71 and 74) levels of uncertainty. However, this phase saw the lowest degree of expert agreement across all assessed tasks, with a median IQR of 2.5.

The nature of uncertainty was the same as observed in the problem formulation phase, with an overall median of 60% for the combined category, and three individual tasks that were either deemed to be epistemic (tasks 70 and 71) or aleatory (task 76; Table 5.5).

Data and variability occurred to the same overall extent as in exposure assessment (median 60%; Table 5.5). However, despite low medians values (of 30%) across all tasks in this phase, experts decided that the extrapolation and model locations were the primary manifestations of uncertainty in group 22, creating stressor-response profiles, both of which had median occurrences of 80%.

Risk characterisation

Risk characterisation contained the highest median level of uncertainty of the four phases, at 5.0 (Figure 5.5d). However, the group of tasks with the highest median level, which concerned assessing the significance of risk levels (5.0; group 27), was lower than the comparable groups in both problem formulation and exposure assessment. The median IQR across all assessed aspects was 2.0, equal to exposure assessment, with a particularly high agreement (IQR of 0.0) seen in task 104.

All 11 tasks were considered to be comprised of both epistemic and aleatory uncertainty, with the combined category yielding a median occurrence of 80% (Table 5.5).

The highest median occurrence rates across risk characterisation (of 60%) were associated with the locations of variability, extrapolation and decision uncertainty, which was the highest seen for the latter location across the three case studies (Table 5.5).

Overall

The median level of uncertainty across all tasks within Case Study 1 was 2.0. The nature-based aspect with the highest median occurrence percentage was the combined category (60%), whilst the location-based uncertainties of data and variability dominated, with overall medians of 60%.

5.5 Results 3: Case Study 2 (particulate matter)

5.5.1 Risk relationship selection

The literature searches, after in-built filtering and relevance-checking, returned 160 peer-reviewed articles, which were further reduced, on the basis of missing information within the articles, yielding a PM evidence base of 61 articles. The PM evidence base was analysed for its risk relationships, the most frequent of which (n=19; Laden *et al.* 2000; Deck *et al.* 2001; Post *et al.* 2001; Goswami *et al.* 2002; Martonen and Schroeter 2003; Sullivan *et al.* 2003; Lai *et al.* 2004; Greene and Morris 2006; Symons *et al.* 2006; Greco *et al.* 2007; Allen *et al.* 2009; Díaz and Dominguez 2009; Jiménez *et al.* 2009; Saldarriaga-Noreña *et al.* 2009; Tainio *et al.* 2010; Betha and Balasubramanian 2011; Boldo *et al.* 2011; Brook *et al.* 2011; Orru *et al.* 2011) was 'potential ambient outdoor $PM_{2.5}$ risk to human health', which was selected as the risk relationship for this case study.

5.5.2 ERA template creation and validation

The generic ERA template, version 3, was populated with relevant information from the 19 peer-reviewed articles, forming the ambient outdoor $PM_{2.5}$ risk to human-health ERA template, version 1 (Appendix H). Validation of this template, through comments (Table 5.6; Supplementary Material D) provided by 8 of the experts in the PM evidence base enabled the creation of the ambient outdoor $PM_{2.5}$ risk to human-health ERA template, version 2 (Appendix J). The experts were based in the sectors of academia (n=4), industry (n=1), and regulation (n=3), and reside in Canada (n=1), Poland (n=1), UK (n=3) and USA (n=3).

Table 5.6 The number of comments received from experts involved in the validation of the $PM_{2.5}$ risk to human-health ERA template, version 1, and the number of changes made to the template (with example changes), organised by ERA phase. The full list of comments is provided in Supplementary Material D.

ERA phase	# comments	# changes	Example change
Problem formulation	25	7	• Included 'primary and secondary' $PM_{2.5}$ in tasks 5, 10 and 11 (3 changes) • Included chemical factors in group 26
Exposure assessment	23	6	• Included time-activity data in task 34, group 8, and in task 55,group 13 (2 changes) • Removed 'carcinogenicity' as a toxicological characteristic (task 39, group 9)
Effects assessment	11	4	• Differentiated between the measures for disease-based assessment endpoints (group 20) • Included new metrics concerning existing toxicity information (task 73, group 18)
Risk characterisation	8	3	• Included ratios, attributable risk, and adjusted life-time risk in task 94 (group 24)

5.5.3 Expert elicitation exercise

Five experts participated in the $PM_{2.5}$ risk to human health elicitation exercise (hereafter referred to as Case Study 2; Table 5.7), with each expert assessing 82 separate ERA-based tasks (26 in problem formulation, 29 in exposure assessment, 16 in effects assessment, and 11 in risk characterisation) for the levels, natures, and locations of associated uncertainty.

Table 5.7 Professional sectors and countries of residence of the experts (n=5) involved in the uncertainty-based elicitation exercise for Case Study 2. Results from the practice exercise are also included, which show the agreement between the experts and the control group with regard to the level (% above or below the control group mean) and nature (% agreement) of uncertainty communicated.

Expert ID	Sector	Country of residence	Level	Nature
1	Regulatory	United States	-	-
2	Regulatory	United States	+15.4%	80%
3	Academia	United Kingdom	-13.0%	80%
4	Regulatory	United Kingdom	+3.8%	60%
5	Academia	United States	+3.8%	80%

The 23 tasks not included in this case study (from the 105 in the generic ERA template, version 3) are shown in Section 5.7. Data relating to the level dimension were treated as non-Gaussian after assessment of the mean, median, and mode values; central tendency and spread were measured using median values and IQRs, respectively.

Problem formulation

Problem formulation had a median level of uncertainty across all of its tasks of 5.0 (Figure 5.6a), 2.0 higher than in the same phase in Case Study 1.

Figure 5.6 The level of uncertainty communicated by the experts (n=5) involved in Case Study 2, across the 82 assessed ERA tasks, organised into the ERA phases of a) problem formulation, b) exposure assessment, c) effects assessment, and d) risk characterisation, and described using median values (red crosses), inter-quartile ranges (boxes), and low-high range values (dashed lines) on a 0 (representing determinism) to 10 (representing total ignorance) scale. The ERA tasks are separated into the groups listed in Table 5.2.

The group of tasks with the highest median level involved choosing assessment and measurement endpoints (7.0; group 4), whilst, as in Case Study 1, experts associated the lowest median levels with the set of tasks aimed at identifying the S-P-R paradigm (3.0; group 3). Experts showed good agreement when assessing the three tasks in group 6, identifying important fate and transport factors, which had IQRs of 1.0 (tasks 26 and 28) and 0.0 (task 27), and overall, with a median IQR across all tasks of 2.0.

The combined epistemic and aleatory category was the dominant aspect of the nature dimension, with an overall occurrence median of 80%, which was 20% higher than the same phase in Case Study 1 (Table 5.8).

Table 5.8 Median occurrence rates (%) for the individual natures and locations of uncertainty provided by experts (n=5) in Case Study 2, organised by ERA phase and showing the highest rates(s; of at least 50%) for the nature (shaded blue) and location (shaded red) of uncertainty associated with each ERA phase (modal values are included for comparison; median occurrence rates on a task-by-task basis are shown in Appendix K).

ERA phase	Nature of uncertainty (%)			Location of uncertainty (%)						
	Epistemic	Aleatory	Combined	Data	Language	System	Variability	Extrapolation	Model	Decision
Problem formulation median	20	0	80	80	20	80	80	20	80	20
Problem formulation mode	20	0	80	80	20	80	80	20	80	20
Exposure assessment median	0	0	80	80	20	40	80	40	80	20
Exposure assessment mode	0	0	80	80	0	40	80	40	80	20
Effects assessment median	0	0	100	100	20	60	100	60	80	20
Effects assessment mode	0	0	100	100	20	80	100	60	80	20
Risk characterisation median	0	0	100	80	20	40	80	80	80	40
Risk characterisation mode	0	0	100	80	20	40	80	60	80	20
Overall median	0	0	80	80	20	70	80	60	80	20
Overall mode	0	0	80	80	20	40	80	60	80	20

There were eight separate tasks for which the combined category had median values of 100%. Only in task 31 was a different option preferred above this category, namely epistemic.

Along with the data and variability locations, which featured heavily in this phase in Case Study 1, system and model uncertainty had median occurrence rates of 80% (Table 5.8). This is the highest median value associated with system uncertainty in any of the case studies. Task 24, determining the relative importance of endpoints to each other, contained a median rate of 100% across all seven locations, indicating that the associated uncertainty manifests in more locations than in any other task in this or either of the other case studies.

Exposure assessment

The median level of uncertainty contained within the exposure assessment phase was also 5.0 (Figure 5.6b), and also 2.0 higher than the same phase in Case Study 1. Group 14, evaluating the stressor-receptor co-occurrence, had the highest median level (7.0). The group with the lowest median level (4.0; group 11), which involved collecting information about the stressor's release, also contained the lowest level in this phase in case study, though the value is 3.5 higher here. All individual ERA tasks reported median levels of between 1.5 and 7.0. The median IQR across all tasks in this phase was just 1.3, intimating higher levels of expert agreement than both the same phase of Case Study 1 or any of the other three phases in this case study.

The nature dimension was again dominated by the combined category, with a phase median of 80%, 20% higher than the same phase in Case Study 1 (Table 5.8). However, two individual tasks (58 and 59) featured elevated occurrence rates (of above 50%) for the aleatory category, the only examples of this across Case Study 2.

Data, variability and model uncertainty all yielded median occurrence rates of 80%, the same as in the preceding phase (Table 5.8). The other four locations returned rates of between 20% and 40%.

Effects assessment

Effects assessment had a higher median level of uncertainty, at 6.0 (Figure 5.6c), than the other three phases in this case study. In particular, group 19, determining the dose of stressor received by the receptor, and group 20, evaluating the assessment endpoints, both contained median levels of 7.0. All individual ERA tasks had median levels of between 4.0 and 8.0, but despite this relatively small range, the median IQR across the tasks in this phase was larger, at 3.0, than all other phases in the first two case studies.

Nine out of the 16 assessed tasks in effects assessment had occurrence rates of 100% for the combined category, which also had an overall median of 100% (Table 5.8). The epistemic category returned median occurrence rates of 0% for all 16 tasks, the only example of this across the case studies.

Data and variability were again the most frequently occurring locations, with median values of 100% (Table 5.8), the same pattern as observed in both of the preceding stages in this case study and in this same phase in Case Study 1.

Risk characterisation

With a median value of 5.0 (Figure 5.6d), lower than the preceding phase, this is the only risk characterisation phase not to contain the highest levels of uncertainty within its respective case study. Despite this, task 101, assessing the confidence in risk levels using experimental evidence, contains the joint-highest median level (8.0) seen in an individual ERA task across the three case studies.

Similar to effects assessment, risk characterisation yielded median occurrence rates of 80% for the combined nature category, with six out of 11 assessed tasks returning median values of 100% (Table 5.8).

Along with the data and variability locations, extrapolation and model uncertainty had median occurrence rates of 80% (Table 5.8). This was the highest median rate for extrapolation uncertainty of any of the phases in the three case studies.

Overall

Across all four phases in Case Study 2, the median level of uncertainty was 5.0, the median occurrence rate for the combined nature category was 100%, and the most frequently occurring locations were data, variability and model uncertainty with median rates of 80%.

5.6 Results 4: Case Study 3 (pesticides)

5.6.1 Risk relationship selection

Due to the vast quantity of risk assessments within the pesticides risk domain, an additional search term of *'water* (in keywords)' was applied to the literature search. This term was chosen in order to maintain parity between the different elements of the environment (Section 4.2.1) that feature in this research, with genetically modified plants and particulate matter closely linked to the elements of land and air, respectively. The literature searches, after in-built filtering and relevance-checking, returned 127 peer-reviewed articles, which were further reduced, on the basis of missing information within the articles, yielding a pesticides evidence base of 49 articles. The pesticides evidence base was analysed for its risk relationships, the most frequent of which (n=5) were 'potential agricultural chemical pesticide risk to surface water macroinvertebrates' and 'potential agricultural chemical pesticide risk to surface water quality'. With dominant parameters for the source and stressor categories (agricultural and chemical, respectively), but not for the receptor category (macroinvertebrates and quality, respectively), several receptor parameters were combined (multiple organisms, algae, crustaceans, and macroinvertebrates), establishing a clearly defined risk relationship (n=13; Cuppen *et al.* 2000; Mastin and Rodgers Jr. 2000; Palma *et al.*2004; van Wijngaarden *et al.* 2004; Wan *et al.* 2006; Schuler and Rand 2008; Siemering *et al.* 2008; van den Brink *et al.* 2009; Vryzas *et al.* 2009; Vryzas *et al.* 2011; Burgert *et al.* 2011; Damásio *et al.* 2011; Guy *et al.* 2011) of 'potential agricultural chemical pesticide risk to surface water organisms', which was selected as the risk relationship for this case study.

5.6.2 ERA template creation and validation

The generic ERA template, version 3, was populated with relevant information from the 13 peer-reviewed articles, forming the agricultural chemical pesticide risk to surface water organisms ERA template, version 1 (Appendix L). Validation of this template, through comments (Table 5.9; Supplementary Material E) provided by 22 of the experts in the PM evidence base enabled the creation of the agricultural chemical pesticide risk to surface water organisms ERA template, version 2 (Appendix M). The experts were based in the sectors of academia (n=15), industry (n=3), and regulation (n=4), and reside in Argentina (n=1), Belgium (n=1), Canada (n=4), China (n=1), Denmark (n=1), France (n=2), Netherlands (n=5), Portugal (n=1), Serbia and Montenegro (n=1), Switzerland (n=1) and USA (n=4).

Table 5.9 The number of comments received from experts involved in the validation of the agricultural chemical pesticide risk to surface water organisms ERA template, version 1, and the number of changes made to the template (with example changes), organised by ERA phase. The full list of comments is provided in Supplementary Material E.

ERA phase	# comments	# changes	Example change
Problem formulation	46	14	• Included 'agricultural spray drift' in tasks 7 and 12 • Included 'biodiversity' as a population-level effect endpoint in group 4
Exposure assessment	37	7	• Maximum residue levels included as existing exposure metrics in group 8 • Included 'partitioning' in task 62, group 15
Effects assessment	20	8	• Four changes made to group 20, assess the effect endpoints, on basis of suggestions both here and in the problem formulation phase • Included species sensitivity distributions as a metric for creating stressor-response profiles (group 22)
Risk characterisation	16	3	• Included potentially affected fraction as a cumulative risk estimation tool (group 24) • Included option to aggregate risk levels for different pathways (group 25)

5.6.3 Expert elicitation exercise

Nine experts participated in the agricultural chemical pesticide risk to surface water organisms elicitation exercise (hereafter referred to as Case Study 3; Table 5.10), with each expert assessing 102 separate ERA-based tasks (31 in problem formulation, 36 in exposure assessment, 21 in effects assessment, and 14 in risk characterisation) for the levels, natures, and locations of associated uncertainty.

Table 5.10 Professional sectors and countries of residence of the experts (n=9) involved in the uncertainty-based elicitation exercise for Case Study 3. Results from the practice exercise are also included, which show the agreement between the experts and the control group with regard to the level (% above or below the control group mean) and nature (% agreement) of uncertainty communicated.

Expert ID	Sector	Nationality	Level	Nature
1	Regulatory	Netherlands	-	-
2	Regulatory	France	-13.0%	60%
3	Academia	Canada	+30.8%	80%
4	Regulatory	Canada	+30.8%	60%
5	Regulatory	Greece	-	-
6	Academia	Switzerland	-	-
7	Industry	United Kingdom	-36.8%	60%
8	Regulatory	Netherlands	-	-
9	Academia	Spain	+23.1%	60%

The 3 tasks not included in this case study (from the 105 in the generic ERA template, version 3) are shown in Section 5.7. Data relating to the level dimension were treated as non-Gaussian after assessment of the mean, median, and mode values; central tendency and spread were measured using median values and IQRs, respectively. The larger dataset size in comparison with the other two case studies also enabled application of the Shapiro-Wilk normality test. The significance values of 16 out of the 102 datasets evaluated (i.e. the ERA tasks) fell below the 0.05 significance threshold, confirming a non-Gaussian dataset.

Problem formulation

The median level of uncertainty across the tasks in problem formulation was 3.0 (Figure 5.7a), the same as for this phase in Case Study 1, but 2.0 lower than in Case Study 2. In another similarity with Case Study 1, the set of tasks aimed at considering the appropriateness of endpoints had the highest associated median level of any group in this phase (6.0; group 5). Experts did not agree to the same extent here compared to the previous two case studies, with a median IQR of 3.0 across all tasks in this phase.

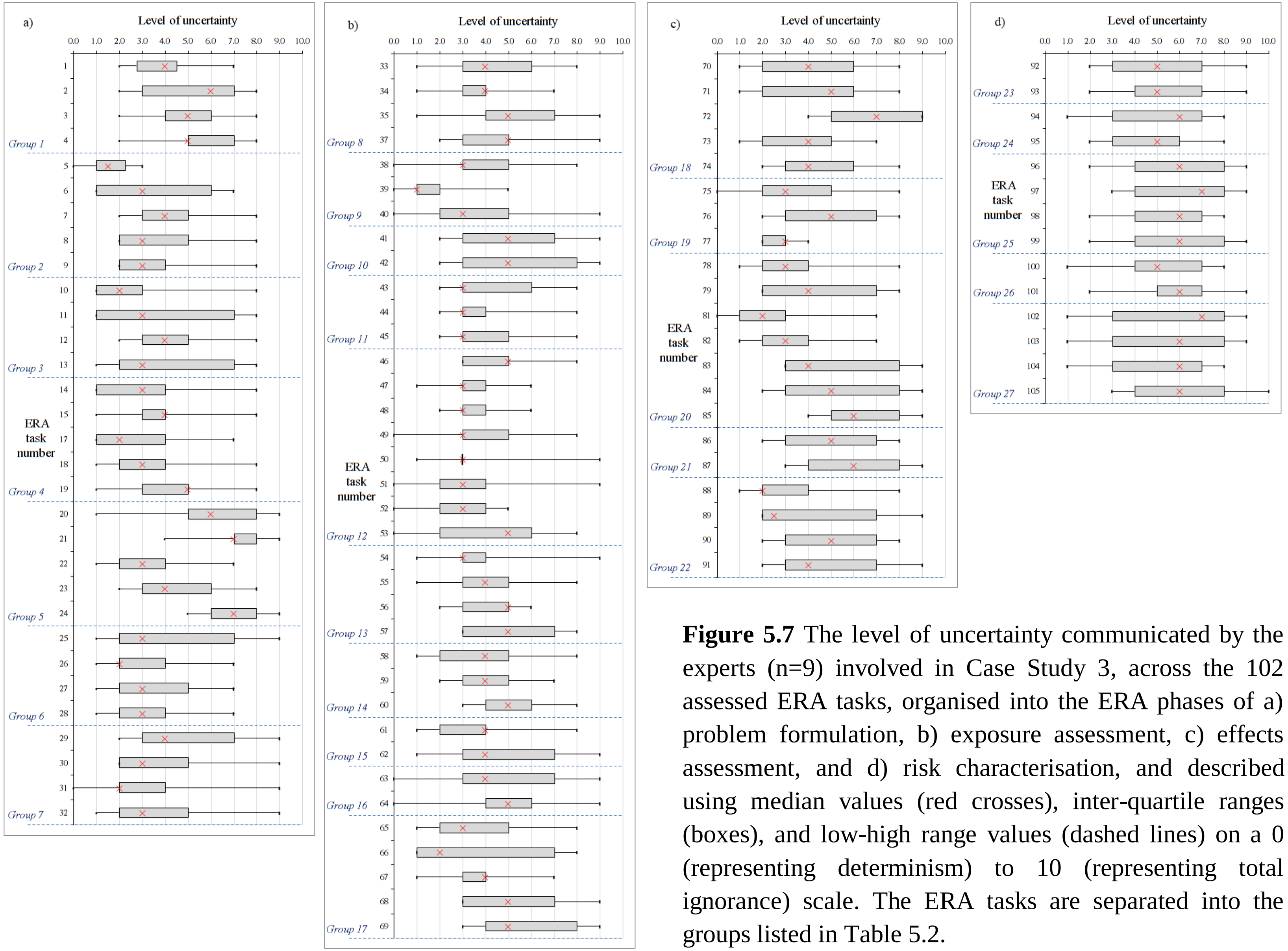


Figure 5.7 The level of uncertainty communicated by the experts (n=9) involved in Case Study 3, across the 102 assessed ERA tasks, organised into the ERA phases of a) problem formulation, b) exposure assessment, c) effects assessment, and d) risk characterisation, and described using median values (red crosses), inter-quartile ranges (boxes), and low-high range values (dashed lines) on a 0 (representing determinism) to 10 (representing total ignorance) scale. The ERA tasks are separated into the groups listed in Table 5.2.

The nature of uncertainty was deemed to consist of the combined epistemic and aleatory category, with a median occurrence rate of 67% across the assessed tasks in this phase (Table 5.11). However, there were six tasks for which no dominant nature was apportioned, due to occurrence rates of below 50%.

Four locations of uncertainty had median occurrence rates of at least 50% in this phase, namely data (67%), system, variability and extrapolation (all 56%; Table 5.11).

Table 5.11 Median occurrence rates (%) for the individual natures and locations of uncertainty provided by experts (n=9) in Case Study 3, organised by ERA phase and showing the highest rates(s; of at least 50%) for the nature (shaded blue) and location (shaded red) of uncertainty associated with each ERA phase (modal values are included for comparison; median occurrence rates on a task-by-task basis are shown in Appendix N).

ERA phase	**Nature of uncertainty (%)**			**Location of uncertainty (%)**						
	Epistemic	**Aleatory**	**Combined**	**Data**	**Language**	**System**	**Variability**	**Extrapolation**	**Model**	**Decision**
Problem formulation median	22	11	67	67	11	56	56	56	44	11
Problem formulation mode	22	0	67	67	0	67	56	56	56	11
Exposure assessment median	33	11	56	67	0	33	50	39	22	11
Exposure assessment mode	33	11	44	56	0	33	56	44	22	11
Effects assessment median	22	11	67	67	0	44	56	33	44	0
Effects assessment mode	11	0	67	67	0	56	67	33	56	0
Risk characterisation median	0	11	89	39	0	56	56	78	78	33
Risk characterisation mode	0	11	89	33	0	44	67	78	78	33
Overall median	22	11	67	67	0	44	56	44	44	11
Overall mode	22	11	67	67	0	33	56	56	44	11

Exposure assessment

The median level of uncertainty in exposure assessment was 4.0 (Figure 5.7b). The set of tasks with the highest median level involved collecting information about the distribution of the stressor (5.0; group 10), whilst several groups contained similarly low median levels (3.0; groups 9, 11 and 12). Experts showed the strongest degree of agreement in this phase compared to the others in this case study, with an overall median IQR of 2.0, but lower than the same phase in Case Study 2.

Half of the 32 assessed ERA tasks in exposure assessment did not have a primary nature of uncertainty associated with them, with the other 16 tasks associated with the combined category (Table 5.11). The phase median for the latter was 56%, the lowest for the combined category in any phase of the three case studies.

Data uncertainty was again the most frequent location based uncertainty, with a median occurrence rate of 67% across the tasks in this phase (Table 5.11). The high level of uncertainty communicated above for group 10 was shown here to manifest primarily through data-based uncertainty, which had a particularly high median rate of 100%.

Effects assessment

Effects assessment here has a lower associated median level of uncertainty than the same phase in the previous two case studies, at 4.0 (Figure 5.7c). Group 22, involving the creation of stressor-response profiles, contained the highest median levels of uncertainty in Case Study 1 (7.0), but here contained the lowest (3.3). Despite the lower levels across this case study, experts disagreed about the values of the individual tasks more than in any other phase across the case studies, with a median IQR of 5.0.

This phase was again dominated by the combined nature category, with a median occurrence rate of 67% (Table 5.11). However, task 70, using available evidence at the beginning of the phase to better constrain the potential effects of the stressor upon the receptor, contained the only example of one of the other two categories occurring more frequently.

As for the previous two phases in this case study, effects assessment consisted primarily of data uncertainty, with a median occurrence rate of 67% (Table 5.11). The uncertainty was also found to exist in the form of variability, but to a lesser extent (56%).

Risk characterisation

Risk characterisation contained the highest median level of uncertainty of the phases in this case study (6.0; Figure 5.7d), and the joint-highest across the case studies. Specifically, groups 25, aggregating risk estimates, and 27, assessing the significance of the risk levels, had the highest associated values, of 6.0. The median levels across the 11 tasks all fell within the range of 5.0 to 7.0, but despite this, the median IQR for this phase was 3.0, larger than for the same phase in Case Studies 1 and 2, highlighting the variation in responses.

All 11 tasks in this phase were deemed to consist of the combined category, with an overall median rate of 89% (Table 5.11).

The locations of extrapolation and model uncertainty were the most frequent here, with median levels of 78% across this phase (Table 5.11), making risk characterisation the only phase in this case study where data uncertainty was not the primary location-based concern.

Overall

The case study median level of uncertainty was 4.0, which places it between the comparable values seen in Case Studies 1 and 2. Experts consistently communicated that the uncertainty seen was both epistemic and aleatory in nature, similar to the other case studies. Whilst data was the joint-highest location-based uncertainty in Case Studies 1 and 2, here, it was the standalone highest with a median rate of 67%.

5.7 Results 5: an uncertainty identification system for environmental risk assessments (UnISERA)

5.7.1 Case study aggregation

The expert responses (n=19) from the three case studies were aggregated (using equal weights; see Section 5.2.4) to form UnISERA, which describes the levels, natures, and locations of uncertainty across 89 separate ERA-based tasks (28 in problem formulation, 32 in exposure assessment, 18 in effects assessment, and 11 in risk characterisation). The ERA tasks not brought forward from the three case studies are shaded grey in Table 5.12.

Table 5.12 The ERA tasks, organised by ERA phase, ERA sub-phase, and ERA task group, included in (denoted by ticks) or excluded from (denoted by crosses) the expert elicitation exercises across the three case studies. ERA tasks excluded from UnISERA are shaded grey.

ERA Task #	ERA Phase	ERA sub-phase	ERA task group	ERA task	Case Study 1	Case Study 2	Case Study 3
1	Problem formulation	Preliminary hazard identification	1. Use available evidence to better constrain…	Potential stressors	✓	✓	✓
2				Potential receptors	✓	✓	✓
3				Potential exposure	✓	✓	✓
4				Potential effects	✓	✓	✓
5			2. Framing the hazard	Frame the 'what'	✓	✓	✓
6				Frame the 'whom'	✓	✓	✓
7				Frame the 'how'	✓	✓	✓
8				Frame the 'where'	✓	✓	✓
9				Frame the 'when'	✓	✓	✓
10		Define the conceptual model	3. Identify the S-P-R paradigm, including…	The source(s)	✓	✓	✓
11				The stressor(s)	✓	✓	✓
12				The pathway(s)	✓	✓	✓
13				The receptor(s)	✓	✓	✓
14			4. Choose assessment and measurement endpoints	Organism: development	✓	✓	✓
15				Organism: behaviour	✓	×	✓
16				Organism: disease	×	✓	×
17				Organism: survival	✓	✓	✓
18				Organism: fecundity	✓	×	✓
19				Population: abundance	✓	×	✓

20				Population: Biodiversity	×	×	✓
21				Ecosystem: PP and NC	×	×	✓
22			5. Consider the appropriateness of the endpoints	Relevance of measures to their endpoints	✓	✓	✓
23				Significance of endpoints to receptor	✓	✓	✓
24				Relative importance of endpoints to each other	✓	✓	✓
25		Form the analysis/work plan	6. Identify the factors controlling fate and transport of the stressor	Biological factors	×	×	✓
26				Chemical factors	×	✓	✓
27				Physical factors	✓	✓	✓
28				Environmental media factors	✓	✓	✓
29			7. Identify data considerations	Gaps in data	✓	✓	✓
30				Types of data required	✓	✓	✓
31				Collection techniques	✓	✓	✓
32				Analysis techniques	✓	✓	✓
33	Exposure assessment	Use available evidence to better constrain…	8. (Use available evidence to better constrain…)	Nature of exposure	✓	✓	✓
34				Exposure levels	✓	✓	✓
35				Model selection	✓	✓	✓
36				Secondary pathways	×	✓	×
37				Prioritisation of data	✓	✓	✓
38		Stressor, exposure media, and receptor information	9. Collect information about the stressor's composition	Biological information	✓	×	✓
39				Chemical information	✓	✓	✓
40				Physical information	✓	✓	✓
41			10. Collect information about the stressor's distribution	Spatial	✓	✓	✓

42			Temporal	✓	✓	✓
43		11. Collect information about the stressor's release	Intensity	✓	✓	✓
44			Probability	✓	✗	✓
45			Quantity	✓	✓	✓
46		12. Collect information about properties affecting fate and transport	Biological	✓	✗	✓
47			Chemical	✗	✓	✓
48			Physical	✓	✓	✓
49			Environmental media: terrestrial	✓	✓	✓
50			Environmental media: biota	✓	✗	✓
51			Environmental media: Sub-terrestrial	✗	✗	✓
52			Environmental media: Atmospheric	✓	✓	✓
53			Environmental media: Aquatic	✗	✗	✓
54		13. Collect information about the receptor	Physical composition	✗	✓	✓
55			Receptor characteristics	✓	✓	✓
56			Spatial distribution	✓	✓	✓
57			Temporal distribution	✓	✓	✓
58	Evaluate stressor-receptor contact	14. Evaluate co-occurrence for…	Spatial overlap	✓	✓	✓
59			Temporal overlap	✓	✓	✓
60			Intensity of overlap	✓	✓	✓
61		15. Evaluate…	Nature of contact	✗	✓	✓
62			Uptake by receptor	✗	✗	✓
63	Integrate multiple LOEs	16. (Integrate multiple LOEs	Semi-quantitative methods	✗	✗	✓

		using…	using…)				
64				Quantitative methods	✗	✓	✓
65		Create the exposure profile(s) using…	17. (Create the exposure profile(s) using…)	Conservative estimates	✓	✓	✓
66				Worst-case estimates	✓	✓	✓
67				Direct monitoring values	✓	✓	✓
68				Stressor-based models	✓	✓	✓
69				Receptor-based models	✓	✓	✓
70	Effects assessment	Use available evidence to better constrain…	18. (Use available evidence to better constrain…)	Nature of effects	✓	✓	✓
71				Direct/indirect effects	✓	✓	✓
72				Secondary stressors	✗	✓	✓
73				Toxicity levels	✓	✓	✓
74				Prioritisation of data	✓	✓	✓
75		Analyse the stressor-response relationship	19. Determine the test dose for the…	Duration	✓	✓	✓
76				Frequency	✓	✓	✓
77				Intensity	✓	✓	✓
78			20. Assess effect endpoints	Organism: development	✓	✓	✓
79				Organism: behaviour	✓	✗	✓
80				Organism: disease	✗	✓	✗
81				Organism: survival	✓	✓	✓
82				Organism: fecundity	✓	✗	✓
83				Population: abundance	✓	✗	✓
84				Population: Biodiversity	✗	✗	✓
85				Ecosystem: PP and NC	✗	✗	✓
86		Integrate multiple LOEs using…	21. (Integrate multiple LOEs using…)	Semi-quantitative methods	✗	✗	✓

87				Quantitative methods	×	✓	✓
88		Create stressor-response profile using…	22. single point methods showing…	Conservative toxicity	✓	✓	✓
89				Extreme toxicity	✓	✓	✓
90				Effects levels	✓	✓	✓
91			distribution methods showing…	Effects levels	✓	✓	✓
92	Risk characterisation	Select relevant profiles…	23. (Select relevant profiles…)	For exposure	✓	✓	✓
93				For effects	✓	✓	✓
94		Estimate and aggregate risk	24. Estimate risk using…	Single-point profiles	✓	✓	✓
95				Cumulative distributions	✓	✓	✓
96			25. Aggregate risk estimates for…	Assessment endpoints	✓	✓	✓
97				Stressors	✓	✓	✓
98				Pathways	×	×	✓
99				Receptors	×	×	✓
100		Evaluate risk levels	26. Assess confidence in the risk levels using…	Empirical evidence	✓	✓	✓
101				Experimental evidence	✓	✓	✓
102			27. Assess the significance of the risk levels using…	Regulatory levels	✓	✓	✓
103				Stakeholder levels	✓	✓	✓
104				Experimental levels	✓	✓	✓
105				Receptor recovery potential	×	×	✓

Problem formulation

Problem formulation contained a median level of uncertainty of 3.4 (Figure 5.8a), compared to 3.0, 3.0, and 5.0 for Case Studies 1, 2 and 3. The group of tasks with the joint-highest median level concerned considering the appropriateness of assessment and measurement endpoints (5.0; group 5; P=0.52), the same observation as in Case Studies 1 and 3. Group 1 (P=0.04), using existing evidence at the beginning of this phase to better constrain certain aspects, shared this median level, largely due to the influence of Case Studies 2 and 3. Group 3, identifying the S-P-R paradigm, contained the lowest median level (3.0; P=0.00), which was also true of all three case studies. Experts agreed to the same extent as in Case Study 3, with a median IQR of 3.0.

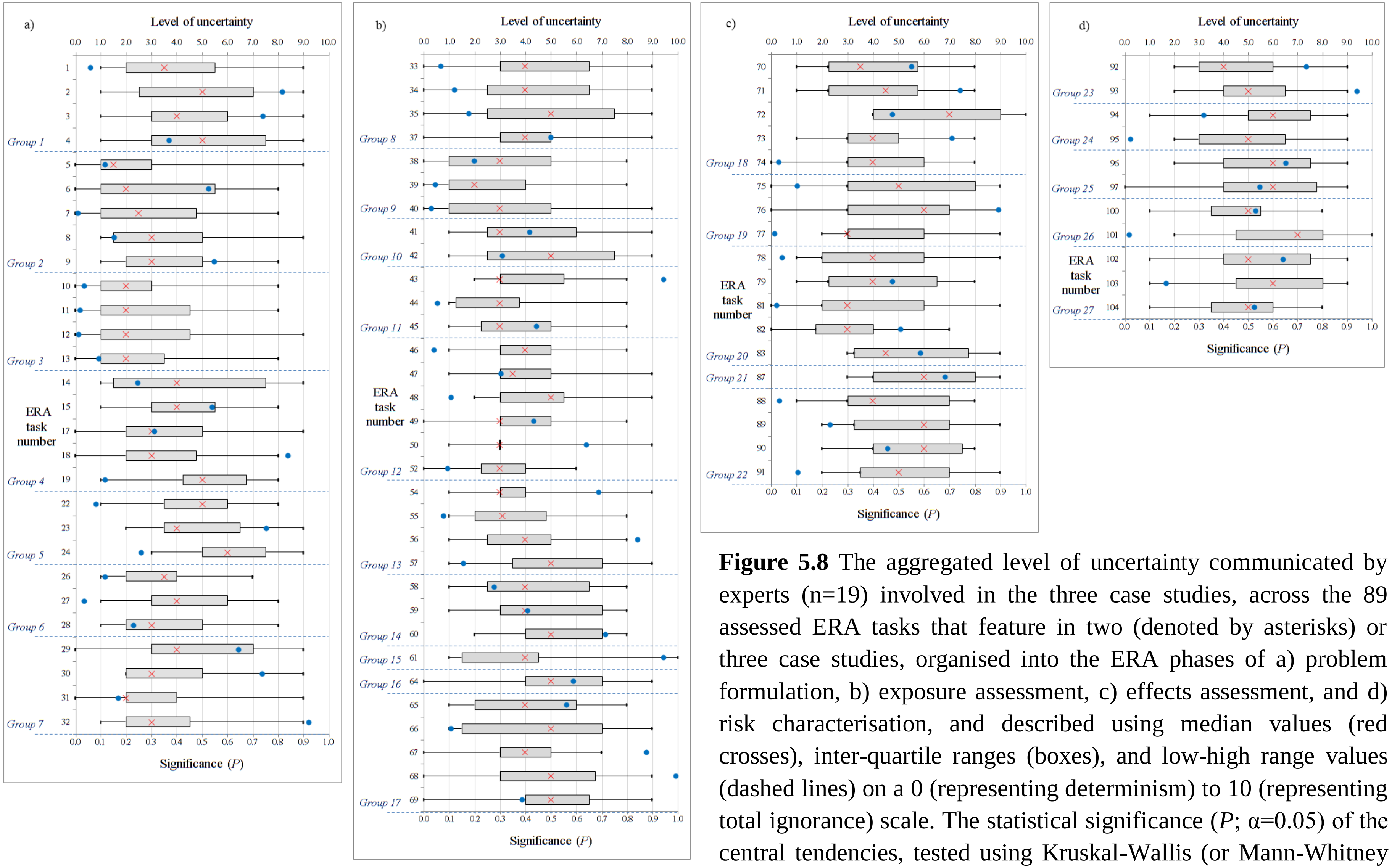


Figure 5.8 The aggregated level of uncertainty communicated by experts (n=19) involved in the three case studies, across the 89 assessed ERA tasks that feature in two (denoted by asterisks) or three case studies, organised into the ERA phases of a) problem formulation, b) exposure assessment, c) effects assessment, and d) risk characterisation, and described using median values (red crosses), inter-quartile ranges (boxes), and low-high range values (dashed lines) on a 0 (representing determinism) to 10 (representing total ignorance) scale. The statistical significance (P; α=0.05) of the central tendencies, tested using Kruskal-Wallis (or Mann-Whitney for ERA tasks with two datasets), are shown (blue circles). The ERA tasks are separated into the groups listed in Table 5.2.

The nature dimension was dominated by the combined epistemic and aleatory category, with a median occurrence rate of 66% (Table 5.13), closely in line with the value reported in Case Study 3. The aleatory option also had its lowest median value across the phase of UnISERA, at just 5%.

Experts returned the highest median occurrence rate (63%) for the location-based uncertainty of data, just ahead of system uncertainty (53%; Table 5.13).

Table 5.13 Median occurrence rates (%) for the individual natures and locations of uncertainty provided by experts (n=19) in UnISERA, organised by ERA phase and showing the highest rates(s; of at least 50%) for the nature (shaded blue) and location (shaded red) of uncertainty associated with each ERA phase (modal values are included for comparison; median occurrence rates on a task-by-task basis are shown in Appendix P).

ERA phase	**Nature of uncertainty (%)**			**Location of uncertainty (%)**						
	Epistemic	**Aleatory**	**Combined**	**Data**	**Language**	**System**	**Variability**	**Extrapolation**	**Model**	**Decision**
Problem formulation median	21	5	66	63	16	53	45	37	45	24
Problem formulation mode	11	5	74	68	16	63	42	26	53	26
Exposure assessment median	21	16	55	63	5	36	58	37	32	15
Exposure assessment mode	21	16	53	63	0	32	68	37	26	11
Effects assessment median	12	8	73	68	5	32	62	49	50	11
Effects assessment mode	11	5	74	74	5	32	32	53	47	5
Risk characterisation median	5	11	84	53	11	47	58	68	58	37
Risk characterisation mode	5	11	84	47	5	53	68	68	68	47
Overall median	19	11	65	62	11	42	55	42	45	21
Overall mode	11	5	74	68	5	53	68	47	53	11

Exposure assessment

Across the tasks in exposure assessment, the median level of uncertainty communicated by experts was 4.0 (Figure 5.8b), the same as in Case Study 3. Group 17, which involved creating the exposure profile(s), contained the highest median level (5.0; P=0.31), which was not seen in this phase in the three case studies. Groups 9 (P=0.00), collecting information about the stressor's composition, and 11 (P=0.31), collecting information about the stressor's release, contained the lowest median level (3.0), correlating with the comparable values in Case Studies 1 and 2, respectively. The degree to which experts agreed on the level values was higher in this phase, with a median IQR of 2.5, than in the other phases in UnISERA.

Whilst the combined nature category again featured most often, with a median rate of 55% (Table 5.13), it did so to a lesser extent than in the other three phases of UnISERA, largely because of the higher influence of the epistemic (21%) and aleatory (16%) options.

Data was the most frequently occurring location-based uncertainty, with a median occurrence of 63%, followed by variability, at 58% (Table 5.13), which was the same pattern observed in Case Studies 1 and 2. Group 14, evaluating the stressor-receptor co-occurrence, was particularly high in variability uncertainty, with a median rate of 74%.

Effects assessment

The median level of uncertainty contained within this phase was 4.3 (Figure 5.8c), higher than the same phase in Case Study 3 and lower than in Case Studies 1 and 2. Excluding group 21 (P=0.68), which contained just one task (number 87), the group with the highest median level (5.5; group 22; P=0.02) concerned creating the stressor-response profile(s), despite the low comparable value in Case Study 3 (of just 3.3). However, effects assessment saw the lowest degree of expert agreement across the phases of UnISERA, with a median IQR of 3.8.

Experts again deemed that the uncertainty associated with the tasks in effects assessment was predominantly epistemic and aleatory in nature, with a median rate for the combined category of 73% (Table 5.13).

Data uncertainty was again the most frequently occurring uncertainty in effects assessment, returning its highest median value across the four phases of UnISERA, at 68% (Table 5.13). Variability also featured to its highest extent, at 62%.

Risk characterisation

Risk characterisation in UnISERA yielded a median level of uncertainty of 5.0 (Figure 5.8d), in line with the same phase in Case Studies 1 and 2. This value was also higher than was seen across the other phases in UnISERA. Groups 25 (P=0.41), aggregating risk estimates, and 26 (P=0.09), assessing the confidence in risk levels, contained the highest level (6.0), whilst group 23, selecting relevant exposure and effects profiles to aggregate, contained the lowest (4.5; P=0.84). The degree of expert agreement across this phase (IQR of 3.5) was slightly better than for effects assessment.

The combined nature category reported its highest phase-by-phase occurrence rate here, with a median value of 84% (Table 5.13), which reflects the observation that the highest (or joint-highest) values for this category were seen in risk characterisation in all three case studies.

Extrapolation uncertainty had the highest associated median occurrence rate (68%), followed by variability and model uncertainties (both 58%; Table 5.13). The extrapolation location was particularly high for the group of tasks associated with estimating risk levels (group 24; 82%), whilst the model location featured most heavily in the subsequent group, which was concerned with aggregating those risk levels (group 25; 68%).

Overall

The median level of uncertainty across all 89 tasks in UnISERA was 4.0, at the lower end of scenario uncertainty. There were no individual tasks across the four phases for which either the epistemic or aleatory natures contained a higher median occurrence rate than the combined category, which had an overall median value of 65%. In terms of the locations in which the uncertainty was manifest across UnISERA, data was the primary concern, with median occurrence rates of at least 50% in 69 out of 89 tasks, followed by variability (57 out of 89), system (35), model (35), extrapolation (29), decision (2), and language (0).

5.7.2 UnISERA

The aggregated results for the individual ERA tasks can be used to describe the order, based on descending values for the level dimension, in which these tasks may be assessed by analysts, depending on priorities. The ten ERA tasks from UnISERA with the highest levels of uncertainty are shown in Table 5.14, along with the specific natures and locations of uncertainty that may be of concern (the full output from UnISERA, organised by descending level of uncertainty, is in Appendix Q, and an electronic version of the results is provided in Supplementary Material F).

Table 5.14 The 10 ERA tasks with the *highest* median levels of uncertainty within UnISERA, with accompanying ranked occurrence rates (with median values of at least 50%) for the nature and locations of uncertainty.

ERA Task #	ERA Phase	ERA sub-phase	ERA task group	ERA task	Level[a]	Nature[b]	Location(s)[c]
72	Effects	Use available evidence to better constrain…	18. (Use available evidence to better constrain…)	Secondary stressors	7.0 (Ig) *P*=0.48	Co	1: Dat; 2: Sys, Mod; 3: Var, Ext;
101	Risk	Evaluate risk levels	26. Assess confidence in the risk levels using…	Experimental evidence	7.0 (Ig) *P*=0.02	Co	1: Ext; 2: Dat, Var;
76	Effects	Analyse the stressor-response relationship	19. Determine the test dose for the…	Frequency	6.0 (Sc) *P*=0.89	Co	1: Var; 2: Mod;
87	Effects	Integrate multiple LOEs using…	21. (Integrate multiple LOEs using…)	Quantitative methods	6.0 (Sc) *P*=0.68	Co	1: Mod; 2: Dat; 3: Var;
96	Risk	Estimate and aggregate risk	25. Aggregate risk estimates for…	Assessment endpoints	6.0 (Sc) *P*=0.65	Co	1: Ext, Mod; 2: Var; 3: Sys;
97	Risk	Estimate and aggregate risk	25. Aggregate risk estimates for…	Stressors	6.0 (Sc) *P*=0.55	Co	1: Mod; 2: Sys, Ext; 3: Var;
90	Effects	Create stressor-response profile using…	22. Single point methods showing…	Effects levels	6.0 (Sc) *P*=0.46	Co	1: Ext, Mod; 2: Dat, Var;
94	Risk	Estimate and aggregate risk	24. Estimate risk using…	Single-point profiles	6.0 (Sc) *P*=0.32	Co	1: Ext, Mod; 2: Var; 3: Dat;
24	Problem	Define the conceptual model	5. Consider the appropriateness of the endpoints	Relative importance of endpoints to each other	6.0 (Sc) *P*=0.26	Co	1: Sys, Mod; 2: Var, Ext; 3: Dat;
89	Effects	Create stressor-response profile using…	22. Single point methods showing…	Extreme toxicity	6.0 (Sc) *P*=0.23	Co	1: Dat, Mod; 2: Var; 3: Ext;

[a] Ig=Recognised ignorance; Sc=Scenario uncertainty. Statistical significance (*P*) is used to rank like values.
[b] Co=Combined.
[c] Dat=Data; Sys=System; Var=Variability; Ext=Extrapolation; Mod=Model. Median occurrence rates are used to rank like values.

Across these ten tasks, risk characterisation contained the highest levels of uncertainty, largely when estimating (task 94), aggregating (tasks 96 and 97) and evaluating (task 101) risk levels, followed by effects assessment, through integrating evidence (task 87) and creating stressor-response profiles (tasks 89 and 90), exposure assessment, through integrating evidence (task 64) and creating exposure profiles (tasks 65 to 69), and problem formulation, though the levels here were comparatively low. The nature of uncertainty for the ERA tasks was exclusively a combination of epistemic and aleatory, for median occurrence rates of at least 50%. Across these ten tasks, effects assessment contained the highest levels of uncertainty, largely when constraining secondary stressors (task 72), determining the frequency with which the receptor is exposed to the stressor (task 76), integrating evidence (task 87) and creating stressor-response profiles (tasks 89 and 90), followed by risk characterisation, largely when estimating risk using single-point profiles (task 94), aggregating risk estimates for assessment endpoints and stressors (tasks 96 and 97) and evaluating risk levels using experimental evidence (task 101). Despite data and variability being the primary concern across all tasks within UnISERA (see Section 5.7.1), model uncertainty is the predominant location in the these 10 tasks, featuring most frequently in seven of them.

5.8 Discussion

5.8.1 Uncertainty across the ERA phases of the case studies

Case Study 1

Different levels, natures, and locations of uncertainty were communicated by the experts across the ERA phases of the three subject domains investigated.

Case Study 1, potential Bt-maize risk to non-target Monarch larvae, contained the lowest median levels of uncertainty for three out of the four ERA phases, with the exception of effects assessment, and the lowest median level overall (i.e. across all tasks and phases). One of the reasons for these low levels is the attention given to this risk relationship. Claims that "transgenic pollen harms Monarch larvae", made by Losey *et al.* (1999), spawned a range of laboratory- and field-based experimental research (Sears *et al.* 2001), which accounts for the relatively large accumulated evidence base in this study (Section 5.4.1) considering the specific nature of the risk relationship, which included just a single stressor and a single receptor. The low levels of uncertainty in this case study were possibly a result of high expert confidence in assessing a specific topic with lots of associated evidence.

Whilst its median levels of uncertainty were relatively low, Case Study 1 contained a median occurrence rate for the aleatory category (of the nature dimension) of 20% across all ERA tasks and phases (Table 5.5), higher than the 0% (Table 5.8) and 11.1% (Table 5.11) seen in Case Studies 2 and 3, respectively. The GMHP-based experts therefore communicated the influence that natural processes and their connected uncertainties can have, even throughout heavily-researched subject domains, where one might expect fewer epistemic uncertainties to exist. It is not surprising then that of the location-based uncertainties across this case study, variability (which is aleatory in nature) is the most dominant, along with data uncertainty. Given its global spotlight (van den Belt 2003), one might assume that this subject-domain carries extensive data records, and that the presence of data uncertainty would be unexpected. However, several authors describe situations of incomplete or unavailable data (Peterson *et al.* 2006; Perry *et al.* 2010) and generally low sample sizes (Rauschen *et al.* 2010).

Case Study 2

The second case study, $PM_{2.5}$ risk to human health, contained the highest median levels of uncertainty across the case studies for three out of the four ERA phases. Effects assessment had the highest levels (median 6.0; Figure 5.6), and was the only example in a case study where risk characterisation did not contain the highest median level of uncertainty. This may be explained by the fact that epidemiological evidence is often required to attribute effects in humans to stressors, as opposed to the direct measurements used in the other case studies (US EPA 2010), which also casts doubt upon the accuracy of concentration-response profiles (Deck *et al.* 2001; Yeh and Small 2002; White *et al.* 2008; Jahn *et al.* 2011). Testing for certain effects in humans is also not always possible, adversely affecting the availability and reliability of the corresponding datasets and further limiting system understanding (Almeida *et al.* 2007; Knol *et al.* 2009b).

These concerns can lead to elevated epistemic uncertainties, supported by epistemic location-based uncertainties occurring more frequently in the effects assessment phase of Case Study 2 (with median proportions of 100% for data, 20% for language, and 60% for system uncertainties; Table 5.8) than in any other phase across the three case studies. However, this may also be a reflection of the fact that this case study contained the highest (or joint-highest) overall occurrence rates for each of the seven location-based uncertainties, not just those belonging to the epistemic set. Specifically, the median occurrence rates for system uncertainty (epistemic) and model uncertainty (combined) were 26% and 36% higher than the nearest corresponding values from the other two case studies, further highlighting the challenges of using epidemiological information in ERAs to enhance knowledge (Schwartz 2002) and as input data for computational and statistical modelling (US EPA 2010).

Case Study 3

Whilst the median levels of uncertainty communicated across the ERA phases for Case Study 3 (agricultural pesticide risk to aquatic surface water organisms) were generally found to lie between the other two case studies, the IQRs were much higher than the corresponding values in Case Studies 1 and 2. A higher IQR is the result of a stretched distribution resulting from disparate values (Manikandan 2011), which equates here to disagreement amongst experts. Case Study 3 contained the broadest risk relationship of the three, into which several potential types of agricultural pesticide, multiple environmental pathways, and a multitude of aquatic species could feasibly fit. Experts base their responses on past experiences and knowledge as well as on the information presented to them (Knol *et al.* 2010); therefore high IQRs may just be the result of disparate experiences. However, they may also result from inherently different attitudes to assessing uncertainty, since the experts involved in this case study provided the most varied responses to the level-based questions during the practice exercise, with an average variance of 26.9% to the control group (Table 5.10), compared with 11.1% for the *Bt*-maize experts of Case Study 1 (Table 5.4), and 12.6% for the PM experts of Case Study 2 (Table 5.7).

The occurrence rates for the natures and locations of uncertainty across the ERA phases followed a similar pattern to the levels. Of note is the location-based uncertainty of extrapolation, which returned high median values in Case Study 3 (Table 5.11), especially during risk characterisation. Several articles from the pesticides subject domain discuss the concerns of extrapolating from, for example, predicted no-effect concentrations (PNECs) and NOAELs in creating risk estimates (primarily quotients; Palma *et al.* 2004; Chèvre *et al.* 2008; Abrantes *et al.* 2010), as well as basing PNECs and NOAELs on questionable data during effects assessment (Uricchio 2004).

When viewed at the ERA phase level, Case Study 1 contained the lowest amounts of uncertainty across the three dimensions, and Case Study 2 the most. For more specific observations to me made, these patterns must be investigated at the ERA task level.

5.8.2 Uncertainty across the ERA tasks of the case studies

Tasks with the lowest levels of uncertainty

Framing uncertainties in the context of specific ERA tasks not only provides more granularity than at the ERA phase level, but allows for more specific and targeted guidance regarding the selection and implementation of uncertainty management techniques.

The ERA tasks with the lowest associated levels of uncertainty (Table 5.15), based on results from the case studies, predominantly reside within the problem formulation phase; just under 50% of the 30 tasks with the lowest levels were found elsewhere.

Table 5.15 The 10 ERA tasks with the *lowest* median levels of uncertainty within the three case-study domains, with accompanying ranked occurrence rates (with median values of at least 50%) for the nature and locations of uncertainty.

ERA Task #	ERA Phase	ERA sub-phase	ERA task group	ERA task	Level[a]	Nature[b]	Location(s)[c]
Bt-maize risk to non-target Monarch butterfly larvae (Case Study 1)							
7	Problem	Preliminary hazard identification	2. Framing the hazard	Frame the 'how'	0.0 (De)	-	-
10	Problem	Define the conceptual model	3. Identify the S-P-R paradigm, including…	The source(s)	0.0 (De)	-	-
12	Problem	Define the conceptual model	3. Identify the S-P-R paradigm, including…	The pathway(s)	0.0 (De)	-	-
11	Problem	Define the conceptual model	3. Identify the S-P-R paradigm, including…	The stressor(s)	1.0 (St)	-	1: Dat;
40	Exposure	Stressor, exposure media, and receptor information	9. Collect information about the stressor's composition	Physical information	1.0 (St)	Co	1: Var; 2: Dat;
13	Problem	Define the conceptual model	3. Identify the S-P-R paradigm, including…	The receptor(s)	1.0 (St)	Ep	1: Lan;
38	Exposure	Stressor, exposure media, and receptor information	9. Collect information about the stressor's composition	Biological information	1.0 (St)	Co	1: Dat, Var;
39	Exposure	Stressor, exposure media, and receptor information	9. Collect information about the stressor's composition	Chemical information	1.0 (St)	Al	1: Var;
5	Problem	Preliminary hazard identification	2. Framing the hazard	Frame the 'what'	1.0 (St)	-	1: Dat, Var;
8	Problem	Preliminary hazard identification	2. Framing the hazard	Frame the 'where'	1.0 (St)	Co	1: Dat, Var;
$PM_{2.5}$ risk to human health (Case Study 2)							
12	Problem	Define the conceptual model	3. Identify the S-P-R paradigm, including…	The pathway(s)	1.0 (St)	-	1: Sys;
55	Exposure	Stressor, exposure media, and receptor information	13. Collect information about the receptor	Receptor characteristics	1.5 (St)	-	1: Ext;
31	Problem	Form the analysis/work plan	7. Identify data considerations	Collection techniques	2.0 (St)	Ep	1: Dat;
13	Problem	Define the conceptual model	3. Identify the S-P-R paradigm, including…	The receptor(s)	2.0 (St)	Co	1: Dat, Sys; 2: Dec;
7	Problem	Preliminary hazard identification	2. Framing the hazard	Frame the 'how'	2.0 (St)	-	-
45	Exposure	Stressor, exposure media, and receptor information	11. Collect information info about the stressor's release	Quantity	2.5 (St)	Co	1: Dat, Var;

41	Exposure	Stressor, exposure media, and receptor information	10. Collect information about the stressor's distribution	Spatial	3.0 (St)	Co	1: Var; 2: Dat, Ext, Mod;
52	Exposure	Stressor, exposure media, and receptor information	12. Collect information about properties affecting fate and transport	Environmental media: Atmospheric	3.0 (St)	Co	1: Dat; 2: Var;
32	Problem	Form the analysis/work plan	7. Identify data considerations	Analysis techniques	3.0 (St)	-	1: Dat;
49	Exposure	Stressor, exposure media, and receptor information	12. Collect information about properties affecting fate and transport	Environmental media: terrestrial	3.0 (St)	-	1: Dat;
Agricultural chemical pesticides risk to surface water organisms (Case Study 3)							
39	Exposure	Stressor, exposure media, and receptor information	9. Collect information about the stressor's composition	Chemical information	1.0 (St)	-	1: Dat;
5	Problem	Preliminary hazard identification	2. Framing the hazard	Frame the 'what'	1.5 (St)	-	-
26	Problem	Form the analysis/work plan	6. Identify the factors controlling fate and transport of the stressor	Chemical factors	2.0 (St)	Co	1: Var; 2: Ext;
10	Problem	Define the conceptual model	3. Identify the S-P-R paradigm, including…	The source(s)	2.0 (St)	-	-
81	Effects	Analyse the stressor-response relationship	20. Assess effect endpoints	Organism: survival	2.0 (St)	Co	1: Dat, Var;
88	Effects	Create stressor-response profile using…	22. Single point methods showing…	Conservative toxicity	2.0 (St)	Co	1: Dat;
17	Problem	Define the conceptual model	4. Choose assessment and measurement endpoints	Organism: survival	2.0 (St)	-	1: Var;
31	Problem	Form the analysis/work plan	7. Identify data considerations	Collection techniques	2.0 (St)	-	1: Dat;
66	Exposure	Create the exposure profile(s) using…	17. (Create the exposure profile(s) using…)	Worst-case estimates	2.0 (St)	-	1: Dat, Ext;
89	Effects	Create stressor-response profile using…	22. Single point methods showing…	Extreme toxicity	2.5 (St)	Co	1: Dat;

[a] De=Determinism; St=Statistical uncertainty; IQRs and high-low ranges are used to rank like values.

[b] Ep=Epistemic; Al=Aleatory; Co=Combined.

[c] Dat=Data; Lan=Language; Sys=System; Var=Variability; Ext=Extrapolation; Mod=Model; Dec=Decision. Median occurrence rates are used to rank like values.

Specifically, the sub-phases of preliminary hazard identification (with five entries) and defining the conceptual model (with eight entries) are areas in which experts had the most confidence. Case Study 1 is the only one in which a median level of determinism (i.e. no uncertainty) is associated with aspects of the ERA, namely tasks 7, 10 and 12, which include identifying the 'how' in hazard framing and the source(s) and pathway(s) in the S-P-R paradigm. Whilst this may indicate an overconfidence in assessing these particular tasks (Morgan and Henrion 1990; Speirs-Bridge *et al.* 2010), it is more likely that a deterministic understanding is communicated here, in the relationship-forming stages of the ERA, because of the unambiguous values considered by those experts, brought about by the narrow risk relationship. The lowest levels of uncertainty seen in Case Studies 2 and 3 are 1.0. Therefore, according to the experts, uncertainty is present in all aspects of these two case studies and should be considered throughout. The natures of these low-level uncertainties cannot be described for 16 of the 30 tasks, since their occurrence rates were below 50% (i.e. they occur less frequently than they occur), similarly for six of the 30 locations. Of the locations that do feature, data uncertainty dominates the three case studies, with variability playing an equal role in Case Study 1. These observations demonstrate the difficulty in providing detailed descriptions of uncertainty for tasks where levels and occurrence rates are low.

Tasks with the highest levels of uncertainty

The tasks with the highest median levels of uncertainty (Table 5.16) are spread fairly evenly across the different ERA phases of the three case studies.

Table 5.16 The 10 ERA tasks with the *highest* median levels of uncertainty within the three case-study domains, with accompanying ranked occurrence rates (with median values of at least 50%) for the nature and locations of uncertainty.

ERA Task #	ERA Phase	ERA sub-phase	ERA task group	ERA task	Level[a]	Nature[b]	Location(s)[c]
			Bt-maize risk to non-target Monarch butterfly larvae (Case Study 1)				
103	Risk	Evaluate risk levels	27. Assess the significance of the risk levels using…	Stakeholder levels	8.0 (Ig)	Co	1: Lan, Sys; 2: Dec;
94	Risk	Estimate and aggregate risk	24. Estimate risk using…	Single-point profiles	7.0 (Ig)		1: Var; Ext; Mod;
88	Effects	Create stressor-response profile using…	22. Single point methods showing…	Conservative toxicity	7.0 (Ig)	Co	1: Ext; Mod; 2: Var;
89	Effects	Create stressor-response profile using…	22. Single point methods showing…	Extreme toxicity	7.0 (Ig)	Co	1: Ext; Mod; 2: Var;
90	Effects	Create stressor-response profile using…	22. Single point methods showing…	Effects levels	7.0 (Ig)	Co	1: Ext; Mod; 2: Var;
27	Problem	Form the analysis/work plan	6. Identify the factors controlling fate and transport of the stressor	Physical factors	6.0 (Sc)	Co	1: Dat; 2: Sys, Var;
22	Problem	Define the conceptual model	5. Consider the appropriateness of the endpoints	Relevance of measures to their endpoints	6.0 (Sc)	Ep	1: Dat, Sys;
93	Risk	Select relevant profiles…	23. (Select relevant profiles…)	For effects	6.0 (Sc)	Co	1: Var, Ext, Dec;
48	Exposure	Stressor, exposure media, and receptor information	12. Collect information about properties affecting fate and transport	Physical	6.0 (Sc)	Co	1: Dat, Mod; 2: Var;
19	Problem	Define the conceptual model	4. Choose assessment and measurement endpoints	Population: abundance	6.0 (Sc)	Co	1: Var; 2: Dat, Ext, Dec;
			$PM_{2.5}$ risk to human health (Case Study 2)				
101	Risk	Evaluate risk levels	26. Assess confidence in the risk levels using…	Experimental evidence	8.0 (Ig)	Co	1: Dat, Var, Ext, Mod;
75	Exposure	Analyse the stressor-response relationship	19. Determine the test dose for the…	Duration	8.0 (Ig)	Co	1: Dat, Var, Mod; 2: Sys; 3: Ext;
91	Effects	Create stressor-response profile using…	22. Distribution methods showing	Effects levels	7.0 (Ig)	Co	1: Dat, Var; 2: Ext;
95	Risk	Estimate and aggregate risk	24.Estimate risk using…	Cumulative distributions	7.0 (Ig)	Co	1: Dat, Var, Ext, Mod; 2: Lan;

57	Exposure	Stressor, exposure media, and receptor information	13. Collect information about the receptor	Temporal distribution	7.0 (Ig)	Co	1: Var; 2: Dat, Ext, Mod;
78	Effects	Analyse the stressor-response relationship	20. Assess effect endpoints	Organism: development	7.0 (Ig)	Co	1: Dat, Var; 2: Sys, Mod;
34	Exposure	Use available evidence to better constrain…	8. (Use available evidence to better constrain…)	Existing exposure levels	7.0 (Ig)	Co	1: Dat, Var; 2: Sys, Mod, Dec
96	Risk	Estimate and aggregate risk	25. Aggregate risk estimates for…	Assessment endpoints	7.0 (Ig)	Co	1: Dat, Var, Ext;
66	Exposure	Create the exposure profile(s) using…	17. (Create the exposure profile(s) using…)	Worst-case estimates	7.0 (Ig)	Co	1: Dat;
80	Effects	Analyse the stressor-response relationship	20. Assess effect endpoints	Organism: disease	7.0 (Ig)	Co	1: Dat, Var; 2: Sys, Ext, Mod;
			Agricultural chemical pesticides risk to surface water organisms (Case Study 3)				
21	Problem	Define the conceptual model	4. Choose assessment and measurement endpoints	Ecosystem: primary prod. and nutrient	7.0 (Ig)	Co	1: Sys; 2: Dat, Mod; 3: Var, Ext;
24	Problem	Define the conceptual model	5. Consider the appropriateness of the endpoints	Relative importance of endpoints to each other	7.0 (Ig)	Co	1: Ext; 2: Sys, Mod; 3: Var;
72	Effects	Use available evidence to better constrain…	18. (Use available evidence to better constrain…)	Secondary stressors	7.0 (Ig)	Co	1: Dat; 2: Sys, Mod;
97	Risk	Estimate and aggregate risk	25. Aggregate risk estimates for…	Stressors	7.0 (Ig)	Co	1: Ext; 2: Sys
102	Risk	Evaluate risk levels	27. Assess the significance of the risk levels using…	Regulatory levels	7.0 (Ig)	Co	1: Ext; 2: Sys;
101	Risk	Evaluate risk levels	26. Assess confidence in the risk levels using…	Experimental evidence	6.0 (Sc)	Co	1: Ext;
85	Effects	Analyse the stressor-response relationship	20. Assess effect endpoints	Ecosystem: PP and NC	6.0 (Sc)	Co	1: Sys; 2: Dat, Ext; 3: Var;
98	Risk	Estimate and aggregate risk	25. Aggregate risk estimates for…	Pathways	6.0 (Sc)	Co	1: Sys, Mod; 2: Ext;
20	Problem	Define the conceptual model	4. Choose assessment and measurement endpoints	Population: Biodiversity	6.0 (Sc)	Co	1: Dat; 2: Sys, Mod; 3: Var, Ext;
87	Effects	Integrate multiple LOEs using…	21. (Integrate multiple LOEs using…)	Quantitative methods	6.0 (Sc)	Co	1: Mod; 2: Dat, Var; 3: Sys, Ext;

[a] Ig=Recognised ignorance; Sc=Scenario uncertainty. IQRs and high-low ranges are used to rank like values.

[b] Ep=Epistemic; Co=Combined.

[c] Dat=Data; Lan=Language; Sys=System; Var=Variability; Ext=Extrapolation; Mod=Model; Dec=Decision. Median occurrence rates are used to rank like values.

A third of these tasks are found in risk characterisation (10 out of 30) despite its smaller size (n=11 tasks) compared to the other phases, with just under a third in effects assessment (9 out of 30). In Case Study 1, the objective of creating the stressor-response profile using single-point methods that show either conservative toxicity (task 88), extreme toxicity (task 89), or effects levels (task 90) all fell within the level of recognised ignorance (i.e. values of between 6.6 and 9.9), with associated epistemic and aleatory uncertainties present, manifesting in the form of extrapolation, model, and variability concerns. Interestingly, the remaining task in this sub-phase, which promotes the use of distributions to show effects levels (task 91), has a much lower level of associated uncertainty (4.0), indicating a preference amongst the participating Bt-maize experts for this probabilistic method over its single-point counterparts. This preference is further evidenced by the existence of task 94 in Table 5.16, which concerns the estimation of risk levels using single-point profiles, a familiar subject in the chemical-based ERA literature (SETAC 1994; US EPA 1999).

The two tasks with the highest level of uncertainty within Case Study 3 appeared in the problem formulation phase, namely the inclusion of the primary production and nutrient cycling endpoint (task 21) and the evaluation of the relative importance of the endpoints to each other (task 24). The former is difficult to measure and is therefore often omitted from ERAs (Barnthouse 2008), and the latter is a reflection of the fact that there are more endpoints in Case Study 3 (n=7) than in Case Study 1 (n=5) or Case Study 2 (n=5; Table 5.12). These early-stage uncertainties are the clear exceptions, with tasks from risk characterisation more commonplace, such as aggregating risk estimates for multiple stressors (task 97) and pathways (task 98) and assessing the significance of these risk estimates using thresholds derived through experimentation (task 101) or regulation (task 102). Case Study 2 reflected this pattern, with all 10 tasks falling outside the problem formulation phase. Across the 30 tasks, the nature of uncertainty was almost exclusively communicated as both epistemic and aleatory combined, and there was representation from all seven locations of uncertainty. These results show that where uncertainty is present, especially at the 'deep' levels (i.e. recognised ignorance) seen in these tasks, it is essential that all aspects of uncertainty are considered and potentially managed (Walker *et al.* 2003; Kandlikar *et al.* 2007; Refsgaard *et al.* 2007; Knol *et al.* 2009a).

Increased levels, natures, and locations of uncertainty allow for informative descriptions and detailed management guidance to be provided, leaving more uncertainty for the analyst to manage. The elicitations from the case studies can be considered extremely useful for analysts based in those specific research domains. However, for the majority grounded in alternative subject matter, a more general description is required.

5.8.3 UnISERA: identifying uncertainty within environmental risk assessments

Identifying uncertainty using the level dimension

In order to direct appropriate resources efficiently, environmental risk analysts may wish to prioritise the aspects of an ERA that have the highest levels of uncertainty associated with them i.e. the phases, sub-phases, and tasks that are 'most uncertain'. In UnISERA, these tasks are primarily associated with the analysis or evaluation of information and evidence rather than its collection or processing. For example, 12 out of the first 20 tasks in UnISERA involve the creation of exposure profiles (tasks 68 and 69), the creation of stressor-response profiles (tasks 89 and 90), and the estimation (task 94), aggregation (tasks 96 and 97), or evaluation (tasks 100, 101, 102, 103, and 104) of risk levels (Appendix Q). From an analysis of nine environmental risk assessments (six ecological and three human), assessed for the magnitude[1], reducibility[2], and quantifiability[3] of uncertainty within, von Stackelberg *et al.* (2008) found that the highest magnitudes of uncertainty were associated with the selection and implementation of profiling metrics during both exposure assessment and effects assessment. The comparable ERA tasks in UnISERA (65 to 69 for exposure metrics and 88 to 91 for effects metrics), some of which are mentioned above, have higher levels of uncertainty associated with them than the majority of the other tasks within the first three ERA phases, but not as high as most tasks within risk characterisation (Figure 5.8). The general trend reported by von Stackelberg *et al.* (2008) is one of increasing uncertainty magnitudes with progression through the four assessment phases; the same pattern was observed in the output from UnISERA, with median uncertainty levels of 3.0 in problem formulation, 4.0 in exposure assessment, 4.3 in effects assessment, and 5.0 in risk characterisation.

[1] The magnitude of uncertainty for aspects of the ERAs was inversely proportional to the completeness of those same aspects; the more certain or complete the input, the lower the magnitude of uncertainty.
[2] Reducibility reflects the effort required to reduce the corresponding magnitude of uncertainty.
[3] Quantifiability is a measure of how easy it is to quantify the corresponding magnitude of uncertainty, depending on its nature.

The concept of three dimensions of uncertainty was first introduced a decade or so ago (Walker *et al.* 2003), and a few expert assessments that apply it to features of risk-based systems (not specifically ERAs) have since been conducted (Krayer von Krauss *et al.* 2004, Gillund *et al.* 2008, Krayer von Krauss *et al.* 2008). Whilst these studies do not extend to the precise locations in which the uncertainty may manifest, they do, to varying extents, involve expert assessments of the levels and natures. After summing and averaging the assessed features (and in one case inverting the scale), the overall levels of uncertainty communicated by the experts across these studies were 3.7 (Krayer von Krauss *et al.* 2004), 4.0 (Krayer von Krauss *et al.* 2008), and 5.1 (Gillund *et al.* 2008), all within the range of scenario uncertainty. The overall median level of uncertainty in UnISERA was 4.1, also firmly within scenario uncertainty.

The natures of uncertainty associated with the 'most uncertain' tasks within UnISERA were exclusively both epistemic and aleatory combined, which indeed extends to all of the tasks in UnISERA for which the nature can be ascribed with confidence (where experts agree to a minimum level of 50%). For reference, only one study (which was used as the basis for the practice exercise in the case study elicitation exercises; Gillund *et al.* 2008) permitted experts to select both epistemic and aleatory in tandem, which was the favoured option the vast majority of the time (in four out of the five tasks assessed). The primary location-based concerns associated with the 'most uncertain' tasks in UnISERA were model and extrapolation uncertainties, which often occurred in tandem. This is probably a consequence of the fact that numerical and statistical models are frequently used in exposure and effects assessments, often with a purpose of extrapolating across species and scales (Forbes *et al.* 2001), with their output a key constituent of exposure and stressor-response profiles (Perry *et al*. 2010). Similarly, model output is often utilised in the estimation and aggregation of risk levels (as evidenced by the existence of mainstream software such as @RISK and Crystal Ball), the evaluation of which can be subject to extrapolation from existing confidence, tolerability, and toxicity thresholds (US EPA 1998; Defra 2011).

As well as identifying the areas of an ERA that harbour the highest levels of uncertainty, it is important to acknowledge the inverse, the 'least uncertain' ERA tasks, since all levels of uncertainty should be managed (van der Sluijs *et al.* 2004; Refsgaard *et al.* 2007). Von Stackelberg *et al.* (2008) found that the lowest magnitudes of uncertainty resided in the problem formulation phase, with which this research agrees (Figure 5.8), and were specifically associated with identifying the source(s; analogous to ERA tasks 1, 5, 10, and

11), pathway(s; ERA tasks 3 and 12), receptor(s; ERA tasks 2, 6, and 13), and suitable assessment and measurement endpoints (ERA tasks 14, 15, 17, 18, and 19), although for this latter task the levels within UnISERA may be described as medium rather than low. The primary location-based uncertainty associated with these tasks in UnISERA was data, which was also connected to other tasks within the sub-phases of preliminary hazard identification and defining the conceptual model (Appendix Q). Data uncertainty occurs frequently in the problem formulation phase, which emphasises the importance of basing initial ERA tasks on reliable datasets, to ensure the adequacy of those tasks as well as the subsequent phases that explore the defined relationships (Wolt *et al.* 2010). The corresponding natures often have median occurrence rates of below 50% (indicating that the potential uncertainty occurred less frequently than it occurred) or invalid nature-location combinations, such as a combined epistemic and aleatory nature being paired with locations of data and system, which are wholly epistemic. This occurs in situations where either one or two locations of uncertainty are dominant (e.g. task 33; Appendix Q) or where several locations are close to the 50% threshold (e.g. task 17; Appendix Q). In such situations of disagreement, appropriate values for the nature should be assigned on the basis of the location(s) of uncertainty.

Risk characterisation contained the highest levels of uncertainty, largely when estimating (tasks 94 and 95), aggregating (tasks 96 to 99) and evaluating (tasks 100 to 104) risk levels, followed by effects assessment, through integrating evidence (task 87) and creating stressor-response profiles (tasks 88 to 91), exposure assessment, through integrating evidence (task 64) and creating exposure profiles (tasks 65 to 69), and problem formulation, though the levels here are comparatively low (Figure 5.8). It should be noted that of these tasks, the comparisons across central tendencies were most significant for numbers 67 (P=0.88; α=0.05) and 68 (P=0.99), with the others no higher than P=0.68. The 20 most statistically relevant tasks within UnISERA (ranging from P=0.99 to P=0.65) were spread evenly, in terms of frequency, across the four phases, with only minor bunching around some numerically adjacent tasks. This indicates that experts from the three case studies agreed on the levels of uncertainty for individual tasks rather than for sets of like tasks within groups or sub-phases. This observation highlights the importance of describing uncertainty within ERAs in as much granularity as possible, that is, on a task-by-task basis.

Identifying uncertainty using the nature and location dimensions

Whilst uncertainty can be approached on a task-by-task or phase-by-phase basis, which effectively prioritises the level dimension above the others, analysts can also implement management options based on the dominant natures and locations of uncertainty. This kind of approach may be useful when resources are limited (e.g. time, money, access to a range of management techniques) or in situations where specific uncertainties are the required focus. The natures of the uncertainties are all a combination of epistemic and aleatory, with no variation from this pattern.

Data uncertainty was the dominant location-based uncertainty within UnISERA, with median occurrence rates of at least 50% in all four phases, 13 out of 15 sub-phases, and 69 out of 89 tasks. The highest rates were seen in the sub-phases of preliminary hazard identification in problem formulation, and collecting stressor, exposure, and receptor information in exposure assessment, both of which are reliant upon data. Besides being data-driven, the problem formulation phase relies on the implementation of system knowledge, and as such can be prone to more system-based uncertainty than other parts of an assessment (Raybould 2006; Wolt *et al.* 2010). In UnISERA, of the ERA tasks with which system uncertainty was most heavily associated, eight of the first 11 were from problem formulation, specifically the sub-phases of preliminary hazard identification and defining the conceptual model. This latter aspect is also susceptible, according to the experts, to model uncertainty, though not as much as the later stages of exposure and effects assessment and risk characterisation, as discussed earlier, which accounted for 12 of the first 14 tasks in which model uncertainty featured most heavily. Another location that impacted risk characterisation was language uncertainty, specifically associated with evaluating the significance of a risk using regulatory and stakeholder levels, synonymous of the potential difficulty in communicating with and drawing information from these groups (Darbra *et al.* 2008). Generally though, experts believed language uncertainty was of little other concern, perhaps not surprising given the relatively small amount of attention attributed to it in the uncertainty and environmental risk literature (Regan *et al.* 2002; Ascough *et al.* 2008). Conversely, variability was the second most frequently occurring uncertainty in UnISERA, behind data. Due to the character of ERAs, natural and human variability can manifest throughout (Huijbregts *et al.* 2001). According to the results from UnISERA, specific attention should be paid to the variability in evaluating the stressor-receptor contact (e.g. spatial, temporal, and intensity of overlap) in exposure assessment, analysing the stressor-response relationship (e.g. effect endpoints) in effects assessment, and selecting relevant exposure and effects profiles for use in risk characterisation. The other location within the aleatory category, extrapolation, is of biggest concern during risk characterisation (discussed earlier), which is the only example of a location-based concern occurring more frequently in an ERA phase than data uncertainty. Finally, decision uncertainty, although not generally a large concern according to the experts, did manifest in individual ERA tasks, most notably in problem formulation (e.g. considering the relative importance of assessment endpoints to each other; task 24) and risk characterisation (e.g. deciding which assessment endpoints to aggregate into final risk levels; task 96). Decision uncertainty may be more influential in post-ERA risk management activities.

5.8.4 Potentially influential methodological aspects

Methodological design

The method of creating and validating system maps to be used as the basis for expert engagement, predominantly to elicit views about risks and uncertainties, is gaining in popularity within the risk community (Krayer von Krauss *et al.* 2004; Gillund *et al.* 2008; Krayer von Krauss *et al.* 2008; Ravnum *et al.* 2012; Smita *et al.* 2012; Zimmer *et al.* 2012).

The method involved the validation of ERA templates, one generic and three domain-specific. Alterations were made to a template where two or more experts provided the same suggestion. The willingness of experts to participate dictated the responses received, as with any form of expert engagement. In order to minimise the amount of alterations required every effort was made to ensure that the templates were as complete and correct as possible, using amassed evidence bases, before being distributed to experts for validation (see Sections 5.2.2 and 5.2.3). However, in sourcing reference materials there is always the chance of an influential source being overlooked. The majority of suggestions corresponded to domain-specific terminology (Supplementary Materials A-E), and across the templates the number of alterations made was small (see Tables 5.1, 5.3, 5.6 and 5.9). The ERA templates used in Case Studies 1 and 3 were altered the least and most, respectively, probably owing to the specific nature of the risk relationship in Case Study 1, which was broader in Case Study 3.

The three risk domains were selected against outlined criteria (Section 5.3), but also involved an element of subjectivity on the part of the researcher. Had different domains been chosen, UnISERA would contain different values. This statement also holds for the specific risk relationships within each domain, though these were selected through having the most associated published material, and were, to a large extent, independent. The question of whether the domains and risk-relationships are too diverse, or not diverse enough, for UnISERA to be representative of uncertainty within ERAs remains (see Chapter 6).

UnISERA is based on aggregated results of expert elicitations. The scope and format of the elicitation can be greatly influenced by the resources available to the researcher (Knol *et al.* 2010). In this case, it was beneficial to conduct remote elicitations. Remotely-executed elicitations have many benefits over face-to-face formats, including that they are far less expensive, their content and structure can be standardised more easily, and experts can complete them at their leisure (US EPA 2009; Knol *et al.* 2010). With the adoption of the remote-elicitation format, it followed that experts be approached as individuals rather than as groups. In this way, the views of experts were considered to be theirs alone, and not that of a collective, as is the case with Delphi methods, for example. The types of experts included was also a consideration.

Three types of experts are noted: generalists, who have substantial knowledge in a discipline connected to the elicitation, and a solid understanding of the elicitation context itself; subject-matter experts, who hold detailed knowledge in the subject of the elicitation and are considered by their peers as an authority in the field; and normative experts, who have knowledge or experience that can aid the elicitation process, such as statistical or psychoanalytical skills (Kotra *et al.* 1996; Loveridge 2004; Knol *et al.* 2010). Experts participating in the elicitation process were required to be subject-matter experts with appropriate levels of domain-expertise, in order for a range of balanced and valid opinions to be communicated (US EPA 2009).

The phrasing of questions and the language used can significantly affect the responses of experts (Payne 1951; Meyer and Brooker 1991). To avoid this influence, all questions (namely the ERA tasks) within the elicitations were worded in a consistent manner, using the validated common terminology within the generic and relevant domain-specific ERA template.

Suggested participant numbers vary depending on the subject, format and budget of the elicitation, and the availability and willingness of experts to participate (US EPA 2009). The inclusion of a minimum of five or six experts is typically considered sufficient to cover the breadth of scientific opinion on a given topic, with little benefit in including additional experts beyond 12 (Clemen and Winkler 1999; Cooke and Probst 2006). This research therefore aimed for participant numbers of between five and 12 per elicitation in order for results to be considered scientifically representative. Incidentally, the frequencies of experts (n=5 for Case Studies 1 and 2, n=9 for Case Study 3) were not capped through choice; the specific nature of two of the case studies in particular (1 and 2) meant that the relevant expert-base from which participants could come was already relatively restricted.

In order to achieve its aim, UnISERA was required to provide definitive levels, natures, and locations of uncertainty across the components of ERAs. Therefore, aggregation into single representative values was required. There is no firm agreement about when aggregation is required and how it should be performed (Keith 1996; Knol *et al.* 2010). There is certainly merit in analysing and communicating the disparate responses of experts. However, even where high levels of value diversity exist, aggregating into single estimates enables comparisons and more effective utilisation of the responses (Knol *et al.* 2010). In some instances, aggregation may not add anything to the reporting of the results (US EPA 2009), but in general, the decision to aggregate is case-specific and depends on the objectives of the elicitation (Krayer von Krauss *et al.* 2004). Numerous aggregation methods exist, including weighted averages (Knol *et al.* 2010), behavioural methods (e.g. Delphi technique), and Bayesian approaches (Clemen and Winkler 1999). Here, equal weights methods were used to aggregate across the 19 sets of results, largely due to the different quantities of experts involved in each case study.

The statistical treatment of collected data can influence its analysis. Typically, tests such as Shapiro-Wilk or Kolmogorov-Smirnov are used to determine whether a distribution can be classified as normal (Gaussian), and by extension as parametric or non-parametric. However, these tests can be unreliable when performed on small datasets (of less than five and seven values for Kolmogorov-Smirnov and Shapiro-Wilk, respectively), since outliers can easily skew otherwise consistently-distributed data. Furthermore, they are not designed for datasets with high frequencies of duplicate values, as seen here. In such circumstances, a more appropriate method, adopted in this research, is to compare the mean, median, and mode, where consistently similar values for each denote normally-distributed data (McCluskey and Lalkhen 2007). Suitable parametric or non-parametric tests can then be applied on the basis of these observations.

It was decided that UnISERA would include the ERA tasks that featured in at least two of the three case studies. Therefore, some of the tasks in UnISERA were based on the aggregated results of two case studies, but the majority on three (See Table 5.12). A legitimate query might follow that for such a system to be truly representative of these three case studies, it must only include like tasks from all three. Doing so would have reduced the size of UnISERA by 16 tasks, six of which involved the selection and evaluation of assessment endpoints common to environmental risk (organism behaviour, organism fecundity, and population abundance). Increasing the number of case studies – which was beyond the scope of this research – and therefore the representativeness of UnISERA, is a principal aim for future work. Whilst the frequencies of experts involved in the elicitations may not be altogether important (provided the minimum and maximum thresholds discussed in Section 5.2.4 are met), the nationalities of experts and the sectors in which they are based are potentially more influential aspects.

The effect of a domain expert's sector and nationality on their responses

It is reasonable to assume that respondents based in industry would be inclined to view the aspects of the respective risk scenarios more favourably (i.e. with more certainty) than those from a regulatory agency, whose job it is to apply sufficient levels of protection (see Section 2.5.6 for information on human variability). However, across all aspects of the three case studies the results do not support this view, with the lowest (or joint lowest) levels of uncertainty communicated by academic experts in 49 out of 89 tasks, compared with 40 and 37 for experts in regulatory and industrial positions, respectively (Figure 5.9a). Further, experts from industry returned the highest (or joint highest) levels of uncertainty on 53 separate occasions, with 44 and 31 for experts from academia and regulation, respectively. The same pattern was observed during the problem formulation and exposure assessment phases, after which academics returned the highest levels, with experts from regulation communicating the lowest in effects assessment, and experts from industry the lowest in risk characterisation. The general pattern across the three case studies was one of academic experts conveying the lowest levels of uncertainty and industrial experts conveying the highest.

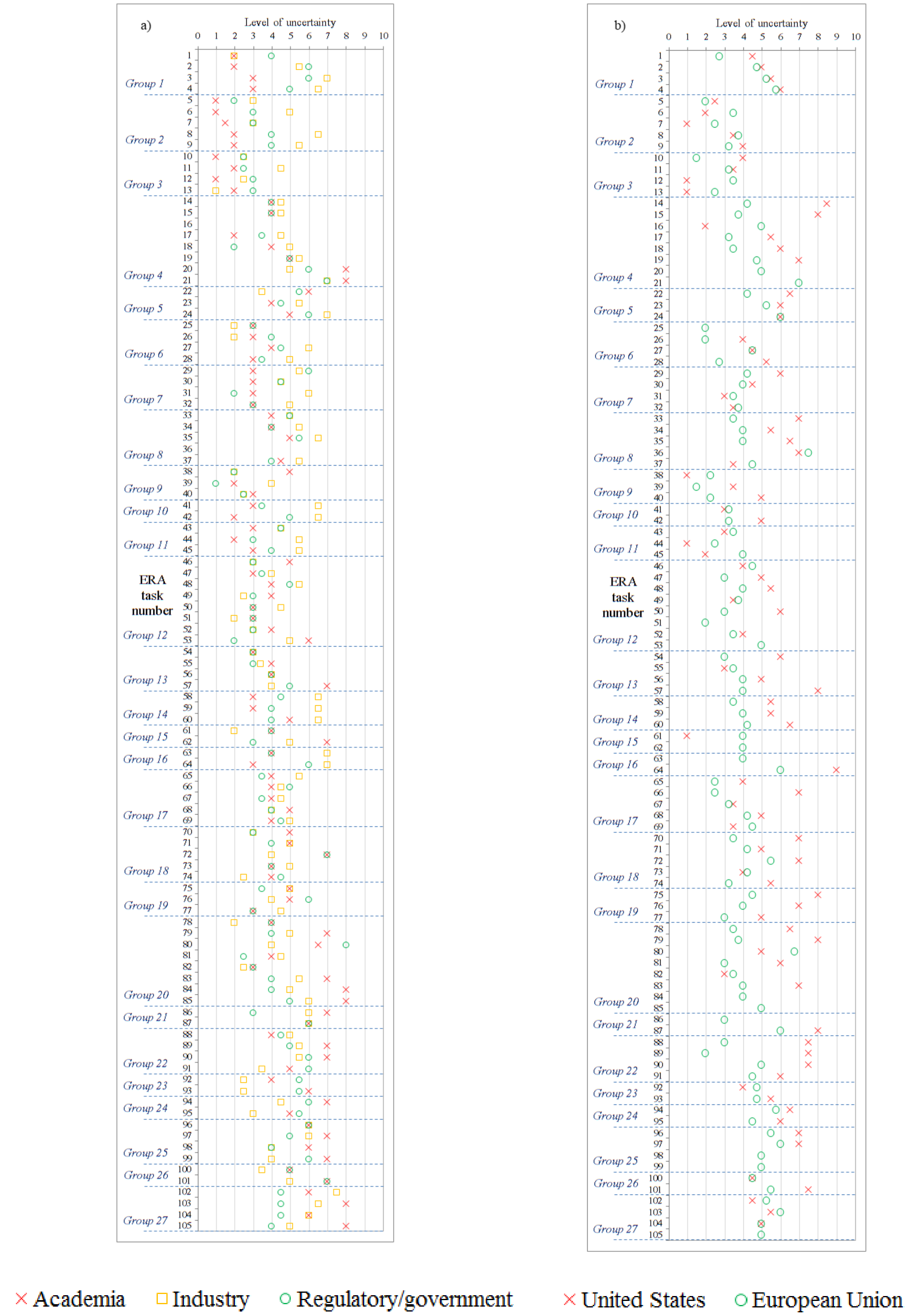


Figure 5.9 The level of uncertainty communicated by experts across the three case studies, organised by task and group, according to a) sector, and b) country of residence. The ERA tasks are separated into the groups listed in Table 5.2.

During the practice exercises experts from academia communicated an average level of uncertainty that was 13% higher than the control group (which itself comprised experts from private and publicly funded research institutes and regulatory agencies; Gillund *et al.* 2008), compared to 9% higher for experts from regulatory positions, and 20% lower for experts from industry. A clear statement regarding the extent to which an expert's sector influences his or her judgement of the level of uncertainty is not possible, based on these two sets of results. Another potentially important aspect to consider is the country in which an expert resides.

In exploring the theme of an expert's nationality as an influence on their responses, the well-established topic of EU versus US approaches to precaution, and by extension uncertainty, is invoked. The strong consensus amongst commentators on the subject is that since the late 1980s, extending to present day, precautionary environmental regulation has shifted from a position of mainstream adoption in the US to that of stringent enforcement in the EU; the so-called 'flip-flop' of regulatory systems (Lofstedt and Vogel 2001). However, a recent comprehensive analysis of risk regulation standards suggests that neither the EU nor the US is more precautionary across the spectrum of environmental risks that these two regions face (Wiener *et al.* 2010).

In broad terms, the research presented here points to US-based experts adopting a more cautious view of the different aspects of the three case studies, with experts from that country proffering higher (or identical) levels of uncertainty in 66 ERA tasks, which was observed on just 30 occasions for experts from EU member countries (Figure 5.9b). Three out of the five participating experts in Case Study 2 were from the USA, which may be a reason for that case study containing a higher overall median level of uncertainty (5.0) than Case Studies 1 (2.0) and 3 (4.0). Of these three experts, two participated in the practice exercise, resulting in an average level of uncertainty that was 8% higher than the control group. This value was in line with the averages of the experts from Case Study 1 (8% higher) and Case Study 2 (7% higher), only one of which resided in the USA.

The inconsistent results across the three case studies and three sets of practice exercises suggest that neither the sector in which an expert works nor the country in which an expert resides consistently influences the responses that they provide regarding the level of uncertainty. Therefore, it is pertinent to conclude that the content of the elicitations was the primary influence on the responses received, as one would hope. This analysis further demonstrates the importance of investigating uncertainty-driven approaches on a risk-by-risk rather than sector-by-sector or country-by-country basis (Wiener *et al.* 2010).

5.9 Conclusion

A reliable characterisation of potential uncertainties can aid in uncertainty identification during ERAs. However, such typologies can be implemented inconsistently, causing uncertainties to go unidentified. Therefore, there is a requirement for a framework which both comprehensively describes uncertainty and offers specific guidance regarding the identification of uncertainties in ERAs.

This chapter developed a system (UnISERA) based on the aggregated results of 19 structured elicitations across three risk domains (genetically modified higher plants, particulate matter, and agricultural pesticides), in which subject-matter experts validated and assessed the aspects of compiled ERA templates for the three dimensions of uncertainty: level; nature; and location. The output from UnISERA describes the uncertainty associated with 89 distinct tasks across the four phases of an ERA; 28 in problem formulation, 32 in exposure assessment, 18 in effects assessment, and 11 in risk characterisation.

The aggregated elicitations revealed that the risk characterisation phase contained the highest median level of uncertainty of 5.0 (on a scale from deterministic understanding of the uncertainty at 0.0, to total ignorance of the uncertainty at 10.0), which were specifically associated with estimating (tasks 94 and 95), aggregating (tasks 96 to 99) and evaluating (tasks 100 to 104) risk levels. Effects assessment contained the second highest median level of uncertainty at 4.3, where integrating evidence (task 87) and creating stressor-response profiles (tasks 88 to 91) were of concern. Exposure assessment had a median level of uncertainty of 4.0, with the highest levels associated with integrating evidence (task 64) and creating exposure profiles (tasks 65 to 69). Problem formulation returned the lowest median level of uncertainty at 3.0, with the highest values related to choosing (task 19) and considering the appropriateness (tasks 22 and 24) of assessment endpoints.

The median nature of uncertainty across the 89 ERA tasks was exclusively a simultaneous combination of epistemic and aleatory. Regarding the locations in which uncertainty was manifest, data uncertainty was dominant in problem formulation, exposure assessment and effects assessment, and had median occurrence rates of at least 50% in 69 out of 89 tasks, followed by variability (57 out of 89), system (35), model (35), and extrapolation (29), which was the dominant location during risk characterisation.

The comprehensive description of uncertainty within UnISERA can be reorganised to fit requirements. This will enable end-users to prioritise ERA phases, tasks, and groups of tasks according to either the highest levels of uncertainty, the potential for the uncertainty to be reduced or only quantified, or the associated types of location-based uncertainty. The novel research presented in this chapter provides specific guidance regarding the identification of uncertainties in ERAs, thereby addressing the outlined knowledge gap. Before this guidance can be accepted and applied by risk analysts, validation of the method and results presented in this chapter must be performed.

Chapter 6: Validating the uncertainty identification system for environmental risk assessments

6.1 Introduction

The introduction of a novel system or set of results, based on either empirical or experimental research, requires testing to ensure that observations are reliable and appropriate for their intended use (González and Herrador 2007). This chapter presents the validation of the methodological approach, output, and insight associated with UnISERA (see Chapter 5), against a distinct ERA domain.

6.2 Method

6.2.1 Overview

The Validation Case Study followed the same methodological approach as the three case studies in Chapter 5, including the use of the generic ERA template (Section 5.2.2), the selection of a risk-relationship and creation of a domain-specific template (Section 5.2.3), and the subsequent execution of expert elicitation exercises (Section 5.2.4).

6.2.2 Data analysis and validation criteria

The data collected were the same as for the three case studies in Chapter 5, with the level of uncertainty represented through integer values and the nature and location of uncertainty treated as binary values. The same forms of analysis were employed, with the level of uncertainty evaluated using central tendency, variation from central tendency, and the high-low range of responses, whilst the date associated with the nature and location of uncertainty was converted to occurrence rates (see Section 5.2.5).

In order to validate UnISERA, its contained values (within the three dimensions) were compared to the corresponding values in the Validation Case Study. Agreement was assessed at different scales, namely across the tasks, groups of tasks, phases, and on an overall basis.

For the level of uncertainty, agreement was noted where corresponding median values were within the same uncertainty bracket (see Figure 2.2), either of determinism (at a level of 0.0), statistical (0.1 to 3.3), scenario (3.4 to 6.6), ignorance (6.7 to 9.9), or total ignorance (10.0). To avoid a situation where two data-points could be similar to each other yet in different level categories (e.g. 3.0 and 3.5), agreement was also noted where there was a maximum difference of 1.0 (i.e. 10% of the 0 to 10 level scale) between corresponding values. The variation in central tendencies of the level values across the tasks, groups of tasks, and phases of UnISERA and the validation study, were also evaluated statistically using the Mann-Whitney test (see Section 5.2.5). The difference between the median level values was also assessed at the different scales.

The nature and location dimensions were assessed in a similar manner to the level dimension, with the three separate nature types and seven separate locations in UnISERA evaluated for their proximity to the corresponding values in the Validation Case Study. The proximity range used when comparing corresponding occurrence rates was set at a maximum of ±33.3%, in order to maintain consistency with the level dimension where corresponding values could be within a maximum of 33.3% of each other (i.e. at the extremes of the same level category) and still be in agreement. Again, agreement was assessed at different scales – across the ERA tasks, groups of tasks, and phases – for both individual locations and for all seven locations combined. The relative difference between the median occurrence rates at these different scales was also evaluated for both dimensions.

6.3 Results

6.3.1 Case study selection

The four criteria outlined in Chapter 5 (see Section 5.3) were also adopted here, with one alteration (the final bullet-point below), namely that a suitable validation risk domain is one:

- where ERAs are prevalent throughout;
- which is relevant to UK-based practitioners (i.e. falls within Defra's landscape of risk concerns);
- where environmental uncertainty is present; and

- which contains a smaller amount of associated empirical evidence than the case studies that comprise UnISERA.

The risk-based domain of focus, based on these criteria, was selected as engineered nanomaterials (see Section 2.3.3 for background information).

6.3.2 Engineered nanomaterials ERA template

Risk relationship selection

The literature searches, after in-built filtering and relevance-checking, returned 84 peer-reviewed articles, which were further reduced, on the basis of missing information within the articles, yielding a ENM evidence base of 50 articles. The ENM evidence base was analysed for its risk relationships, though a dominant relationship was not established. Many of the articles did not focus on specific sources, stressors, or receptors, but instead conducted ERAs (or parts of ERAs, such as exposure assessments) either with unspecified or multiple potential options. For this reason, the most frequently occurring aspects within the evidence base were used to create a composite risk relationship, with information being drawn from the appropriate section(s) of the corresponding articles. These most-frequent aspects were identified as consumer-based engineered nanomaterials for the source (n=20; Handy *et al.* 2008; Mueller and Nowack 2008; Madl and Pinkerton 2009; Thomas *et al.* 2009; Gottschalk *et al.* 2010; Musee *et al.* 2010; Aschberger *et al.* 2011; Farkas *et al.* 2011; Johnson *et al.* 2011; Lorenz *et al.* 2011; Musee 2011; Shaw and Handy 2011; Som *et al.* 2011; Thomas *et al.* 2011; Wang *et al.* 2011; Biskos and Schmidt-Ott 2012; Chio *et al.* 2012; Lapresta-Fernández *et al.* 2012; Matranga and Corsi 2012; Olson and Gurian 2012), nano-Ag (nanosilver) for the stressor (n=8; Mueller and Nowack 2008; Gottschalk *et al.* 2010; Aschberger *et al.* 2011; Farkas *et al.* 2011; Lorenz *et al.* 2011; Musee 2011; Chio *et al.* 2012; Lapresta-Fernández *et al.* 2012), and freshwater fish for the receptor (n=16; Zhu *et al.* 2007; Griffitt *et al.* 2008; Handy *et al.* 2008; Zhu *et al.* 2008; Aschberger *et al.* 2011; Chen *et al.* 2011; Farkas *et al.* 2011; Johnson *et al.* 2011; Quik *et al.* 2011; Shaw and Handy 2011; Thomas *et al.* 2011; Wang *et al.* 2011; Chio *et al.* 2012; Eckelman *et al.* 2012; Lapresta-Fernández *et al.* 2012; Matranga and Corsi 2012), yielding a risk relationship of 'potential consumer-based engineered nanomaterials risk to freshwater fish', which had a collective pool of 26 ERAs, or sections of ERAs, from which information was collected.

ERA template creation and validation

The generic ERA template, version 3 (see Section 5.3), was populated with relevant information from the 26 peer-reviewed articles, forming the consumer-based engineered nanomaterials risk to freshwater fish ERA template, version 1 (Appendix R). Validation of this template, through comments (Table 6.1; Supplementary Material G) provided by 9 of the experts in the ENM evidence base enabled the creation of the consumer-based engineered nanomaterials risk to freshwater fish ERA template, version 2 (Appendix S). The experts were from the sectors of academia (n=5), industry (n=1), and regulation/government (n=3), and the countries of Canada (n=1), Germany (n=1), Luxembourg (n=1), Netherlands (n=1), Sweden (n=1), UK (n=1) and USA (n=3).

Table 6.1 The number of comments received from experts involved in the validation of the consumer-based engineered nanomaterials risk to freshwater fish ERA template, version 1, and the number of changes made to the template (with example changes), organised by ERA phase. The full list of comments is provided in Supplementary Material G.

ERA phase	# comments	# changes	Example change
Problem formulation	11	3	• Added 'ionic nano-Ag' to the stressor term (task 11) • Included 'bioaccumulation' as a fate and transport control process (task 25)
Exposure assessment	10	1	• Inserted the word 'contaminated' into task 61 (group 15)
Effects assessment	7	1	• Included benchmark dose as a method for creating effects-based stressor-response profile(s)
Risk characterisation	4	0	No changes made

6.3.3 Nanosilver uncertainty-based expert elicitation

A total of seven experts participated in the consumer-based engineered nano-Ag risk to freshwater fish elicitation exercise (hereafter termed the Validation Case Study; Table 6.2), though for one of the experts (number seven) the results from the practice exercise fell outside the acceptable agreement range for both the level and nature dimension.

Table 6.2 Professional sectors and countries of residence of the experts involved (and the one expert not involved, shaded grey) in the uncertainty-based elicitation exercises for the Validation Case Study. Results from the practice exercise are also included, which show the agreement between the experts and the control group with regard to the level (% above or below the control group mean) and nature (% agreement) of uncertainty communicated.

Expert ID	Sector	Country of residence	Level	Nature
1	Regulatory/Government	Luxembourg	+26.9%	80%
2	Academia	United States	-	-
3	Academia	Netherlands	-	-
4	Regulatory/Government	United Kingdom	-4.0%	60%
5	Regulatory/Government	United States	0.0%	80%
6	Academia	Sweden	-	-
7	Academia	United States	-62%	20%

The remaining six experts each assessed 99 separate ERA-based tasks (30 in problem formulation, 35 in exposure assessment, 20 in effects assessment, and 14 in risk characterisation) for the levels, natures, and locations of associated uncertainty. The assessed tasks, brought forward from the consumer-based engineered nano-Ag risk to freshwater fish ERA template, version 2, are shown in Table 6.3. Level-based data were treated as non-Gaussian after assessment of the mean, median, and mode values; central tendency and spread were measured using median values and IQRs, respectively.

Table 6.3 The ERA tasks, organised by ERA phase, ERA sub-phase, and ERA task group, included in (denoted by ticks) or excluded from (denoted by crosses) the Validation Case Study and UnISERA. ERA tasks included in UnISERA but not in the Validation Case Study (and therefore not validated) are shaded grey (see Section 6.3.4).

ERA Task #	ERA Phase	ERA sub-phase	ERA task group	ERA task	Validation Case Study	UnISERA
1	Problem formulation	Preliminary hazard identification	1. Use available evidence to better constrain…	Potential stressors	✓	✓
2				Potential receptors	✓	✓
3				Potential exposure	✓	✓
4				Potential effects	✓	✓
5			2. Framing the hazard	Frame the 'what'	✓	✓
6				Frame the 'whom'	✓	✓
7				Frame the 'how'	✓	✓
8				Frame the 'where'	✓	✓
9				Frame the 'when'	✓	✓
10		Define the conceptual model	3. Identify the S-P-R paradigm, including…	The source(s)	✓	✓
11				The stressor(s)	✓	✓
12				The pathway(s)	✓	✓
13				The receptor(s)	✓	✓
14			4. Choose assessment and measurement endpoints	Organism: development	✓	✓
15				Organism: behaviour	✓	✓
16				Organism: disease	✗	✗
17				Organism: survival	✓	✓

18				Organism: fecundity	✓	✓
19				Population: abundance	✓	✓
20				Population: Biodiversity	✗	✗
21				Ecosystem: PP and NC	✓	✗
22			5. Consider the appropriateness of the endpoints	Relevance of measures to their endpoints	✓	✓
23				Significance of endpoints to receptor	✓	✓
24				Relative importance of endpoints to each other	✓	✓
25		Form the analysis/work plan	6. Identify the factors controlling fate and transport	Biological factors	✓	✗
26				Chemical factors	✓	✓
27				Physical factors	✓	✓
28				Environmental media factors	✓	✓
29			7. Identify data considerations	Gaps in data	✓	✓
30				Types of data required	✓	✓
31				Collection techniques	✓	✓
32				Analysis techniques	✓	✓
33	Exposure assessment	Use available evidence to better	8. (Use available evidence to better constrain…)	Nature of exposure	✓	✓
34				Exposure levels	✓	✓
35				Model selection	✓	✓
36				Secondary pathways	✓	✗
37				Prioritisation of data	✓	✓
38		Stressor, exposure media, and receptor	9. Collect information about the stressor's composition	Biological information	✗	✓
39				Chemical information	✓	✓

40				Physical information	✓	✓
41			10. Collect information about the stressor's distribution	Spatial	✓	✓
42				Temporal	✓	✓
43			11. Collect information about the stressor's release	Intensity	✓	✓
44				Probability	✓	✓
45				Quantity	✓	✓
46			12. Collect information about properties affecting fate and	Biological	✓	✓
47				Chemical	✓	✓
48				Physical	✓	✓
49				Environmental media: terrestrial	✓	✓
50				Environmental media: biota	✗	✓
51				Environmental media: Sub-terrestrial	✓	✗
52				Environmental media: Atmospheric	✓	✓
53				Environmental media: Aquatic	✓	✗
54			13. Collect information about the receptor	Physical composition	✓	✓
55				Receptor characteristics	✓	✓
56				Spatial distribution	✓	✓
57				Temporal distribution	✓	✓
58		Evaluate stressor-receptor contact	14. Evaluate co-occurrence for…	Spatial overlap	✓	✓
59				Temporal overlap	✓	✓
60				Intensity of overlap	✓	✓

61			15. Evaluate…	Nature of contact	✓	✓
62				Uptake by receptor	✓	✗
63		Integrate multiple LOEs using…	16. (Integrate multiple LOEs using…)	Semi-quantitative methods	✓	✗
64				Quantitative methods	✓	✓
65		Create the exposure profile(s) using…	17. (Create the exposure profile(s) using…)	Conservative estimates	✓	✓
66				Worst-case estimates	✓	✓
67				Direct monitoring values	✓	✓
68				Stressor-based models	✓	✓
69				Receptor-based models	✓	✓
70	Effects assessment	Use available evidence to better	18. (Use available evidence to better constrain…)	Nature of effects	✓	✓
71				Direct/indirect effects	✓	✓
72				Secondary stressors	✓	✓
73				Toxicity levels	✓	✓
74				Prioritisation of data	✓	✓
75		Analyse the stressor-response relationship	19. Determine the test dose for the…	Duration	✓	✓
76				Frequency	✓	✓
77				Intensity	✓	✓
78			20. Assess effect endpoints	Organism: development	✓	✓
79				Organism: behaviour	✓	✓
80				Organism: disease	✗	✗
81				Organism: survival	✓	✓
82				Organism: fecundity	✓	✓
83				Population: abundance	✓	✓
84				Population: Biodiversity	✗	✗

85				Ecosystem: PP and NC	✓	×
86		Integrate multiple LOEs using…	21. (Integrate multiple LOEs using…)	Semi-quantitative methods	✓	×
87				Quantitative methods	✓	✓
88		Create stressor-response profile	22. Single point methods showing…	Conservative toxicity	✓	✓
89				Extreme toxicity	✓	✓
90				Effects levels	✓	✓
91			(Distribution methods showing…)	Effects levels	✓	✓
92	Risk characterisation	Select relevant profiles…	23. (Select relevant profiles…)	For exposure	✓	✓
93				For effects	✓	✓
94		Estimate and aggregate risk	24. Estimate risk using…	Single-point profiles	✓	✓
95				Cumulative distributions	✓	✓
96			25. Aggregate risk estimates for…	Assessment endpoints	✓	✓
97				Stressors	✓	✓
98				Pathways	✓	×
99				Receptors	✓	×
100		Evaluate risk levels	26. Assess confidence in the risk levels using…	Empirical evidence	✓	✓
101				Experimental evidence	✓	✓
102			27. Assess the significance of the risk levels using…	Regulatory levels	✓	✓
103				Stakeholder levels	✓	✓
104				Experimental levels	✓	✓
105				Receptor recovery potential	✓	×

Problem formulation

The median level of uncertainty seen in the problem formulation phase was 4.0 (Figure 6.1a). The set of tasks which aimed to consider the data requirements of the ERA (group 7) had the highest associated median level, of 5.8. This group also contained the individual task with the highest median level (task 31; 8.0), which concerned identifying appropriate data collection techniques. Experts agreed on the values in this phase to a good extent, with a median IQR across all assessed tasks of 1.8.

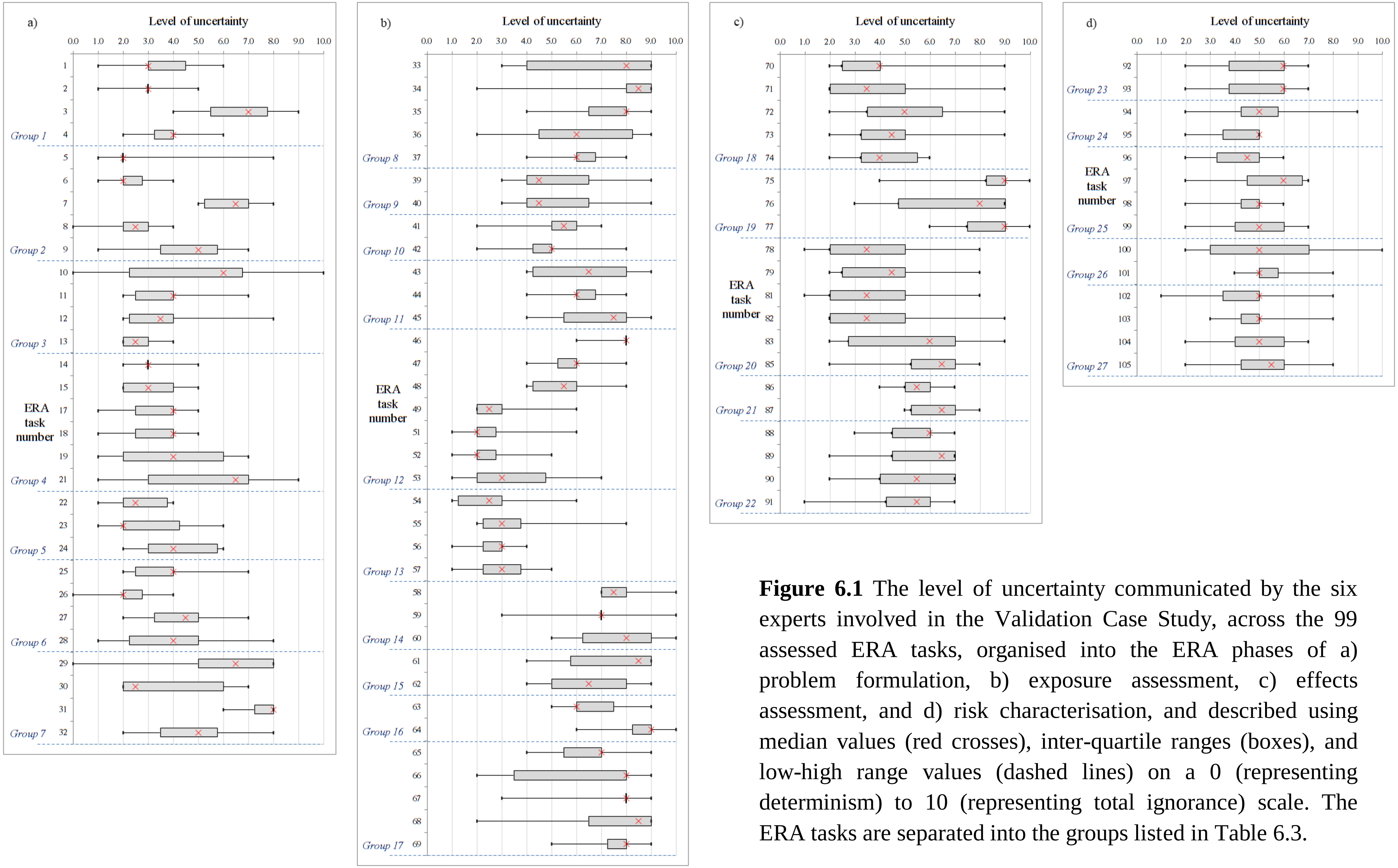


Figure 6.1 The level of uncertainty communicated by the six experts involved in the Validation Case Study, across the 99 assessed ERA tasks, organised into the ERA phases of a) problem formulation, b) exposure assessment, c) effects assessment, and d) risk characterisation, and described using median values (red crosses), inter-quartile ranges (boxes), and low-high range values (dashed lines) on a 0 (representing determinism) to 10 (representing total ignorance) scale. The ERA tasks are separated into the groups listed in Table 6.3.

The combined nature category had a median occurrence rate of 67% in this phase (Table 6.4), with experts deciding that group 4, which involved choosing assessment and measurement endpoints, had a median rate of 100%. The epistemic and aleatory categories had rates of 17% and 0%, respectively, which remained the same for the subsequent three phases of the Validation Case Study.

The data and system locations of uncertainty occurred most frequently with median levels of 67%, followed by variability, at 50% (Table 6.4).

Table 6.4 Median occurrence rates (%) for the individual natures and locations of uncertainty provided by experts (n=6) in the Validation Case Study, organised by ERA phase and showing the highest rates(s; of at least 50%) for the nature (shaded blue) and location (shaded red) of uncertainty associated with each ERA phase (modal values are included for comparison; median occurrence rates on a task-by-task basis are shown in Appendix T).

ERA phase	Nature of uncertainty (%)			Location of uncertainty (%)						
	Epistemic	Aleatory	Combined	Data	Language	System	Variability	Extrapolation	Model	Decision
Problem formulation median	17	0	67	67	17	67	50	33	17	0
Problem formulation mode	0	0	50	67	17	83	50	33	17	0
Exposure assessment median	17	0	67	83	0	50	67	17	17	0
Exposure assessment mode	17	0	83	100	0	50	67	17	17	0
Effects assessment median	17	0	83	83	0	67	67	33	17	8
Effects assessment mode	17	0	83	67	0	67	67	17	17	0
Risk characterisation median	17	0	83	33	17	50	33	50	17	50
Risk characterisation mode	17	0	67	33	0	17	33	33	17	50
Overall median	17	0	67	83	17	50	50	33	17	0
Overall mode	17	0	83	83	0	83	67	17	17	0

Exposure assessment

Exposure assessment contained the highest median level of uncertainty of the four phases in the Validation Case Study, at 6.0 (Figure 6.1b). Group 17, which involved creating the exposure profile(s), was believed by experts to harbour the most significant levels of uncertainty, with a median value of 8.0 across its tasks. Conversely, group 13, which concerned the collection of receptor-based information, had a median level of just 3.0. The individual task with the highest level of uncertainty (9.0) was task 64, integrating multiple lines of evidence using quantitative methods, whilst the lowest level (2.0) was seen in tasks 51 and 52, which required the collection of fate and transport data. The individual tasks therefore spanned the statistical, scenario, and ignorance forms of uncertainty. Despite the large range in responses, this phase saw the highest degree of expert agreement of all phases in the Validation Case Study, with an overall median IQR of just 1.5.

The combined nature category occurred most frequently in exposure assessment, with a median of 67% (Table 6.4). Just one set of tasks, group 9, collecting information about the composition of the stressor, was considered to be primarily epistemic (67%).

The location of data uncertainty recorded a median rate of 83% across this phase (Table 6.4), with values of 100% for groups 11, collecting information about the stressor's release, 15, evaluating the stressor-receptor contact, and 16, integrating multiple lines of evidence. There were six individual tasks (49, 51, 52, 54, 56 and 57) with associated medians 0.0%, all of which had high values for variability and low values (or 0%) for the remaining six categories.

Effects assessment

The median level of uncertainty for this phase was 5.5 (Figure 6.1c). The group of tasks which involved determining the duration, frequency and intensity of the stressor dose received by the receptor (group 19) had the highest associated median level (9.0) of any group in the Validation Case Study. However, effects assessment as a whole saw the lowest degree of expert agreement in the case study, with a median IQR of 2.4.

As with the other three phases in the Validation Case Study, the combined nature category occurred most frequently, here with a median value of 83% (Table 6.4). However, group 19, which was concerned with determining the dose of the stressor received by the receptor, was largely epistemic (67%).

Whilst the data location returned the highest median occurrence rate in effects assessment (83%; Table 6.4), including a rate of 100% for group 19, the system and variability categories were also highly preferred (both 67%). Despite its low overall value (of 33%), extrapolation uncertainty featured heavily in group 20, the assessment of effect endpoints, with a median rate of 75%.

Risk characterisation

The median level of uncertainty across the 14 tasks in risk characterisation was 5.0 (Figure 6.1d). The individual tasks contained median levels of between 4.5 and 6.0, all in the range of scenario uncertainty. Similarly, the five groups in this phase (groups 23 to 27) all had median levels of between 5.0 and 6.0.

The nature dimension returned the same set of median values across its three categories as in the previous phase (Table 6.4).

There were three locations of uncertainty with median occurrence rates of 50% in this phase, namely system, extrapolation and decision, meaning that risk characterisation was the only phase not to be dominated by data uncertainty (Table 6.4). Two tasks, assessing the significance of risk estimates using thresholds defined by regulators (task 102) and stakeholders (task 103), also contained the highest median rates of language (50%) and decision (67%) uncertainty seen across the 99 tasks assessed.

Overall

The median level of uncertainty across all 99 tasks in the Validation Case Study was 5.0. Experts shared a high degree of agreement in this case study, with a median IQR of 1.8. The combined nature category was dominant, with a median occurrence rate of 67%, and the location-based source of data uncertainty was the primary concern for experts (83%), followed by system and variability (both 50%).

6.3.4 Validation of UnISERA

A total of 87 out of 89 ERA tasks in UnISERA were validated against the corresponding tasks in the Validation Case Study. Two tasks (numbers 38 and 50) could not be validated since they did not feature in the Validation Case Study.

Problem formulation

The median level of uncertainty in the Validation Case Study was 4.0, compared with 3.4 for UnISERA. The level of uncertainty (Figure 6.2a) was in agreement for 16 out of the 28 tasks in problem formulation (57% agreement). The four tasks in group 4, choosing assessment and measurement endpoints, all agreed with the comparable tasks in UnISERA, yielding an agreement level of 100% (P=0.25). Task 6, framing the 'whom' of hazard identification, also shared the same median value (2.0; P=0.89) with the same task in UnISERA, meaning that, according to this measure of central tendency, a perfect agreement was observed. When statistically measuring similarity through comparison of non-parametric central tendencies, group 1 (P=0.56), using existing evidence at the start of the phase to constrain certain aspects, and task 27 (P=0.92), identifying the physical factors that control fate and transport processes, recorded the highest significance values.

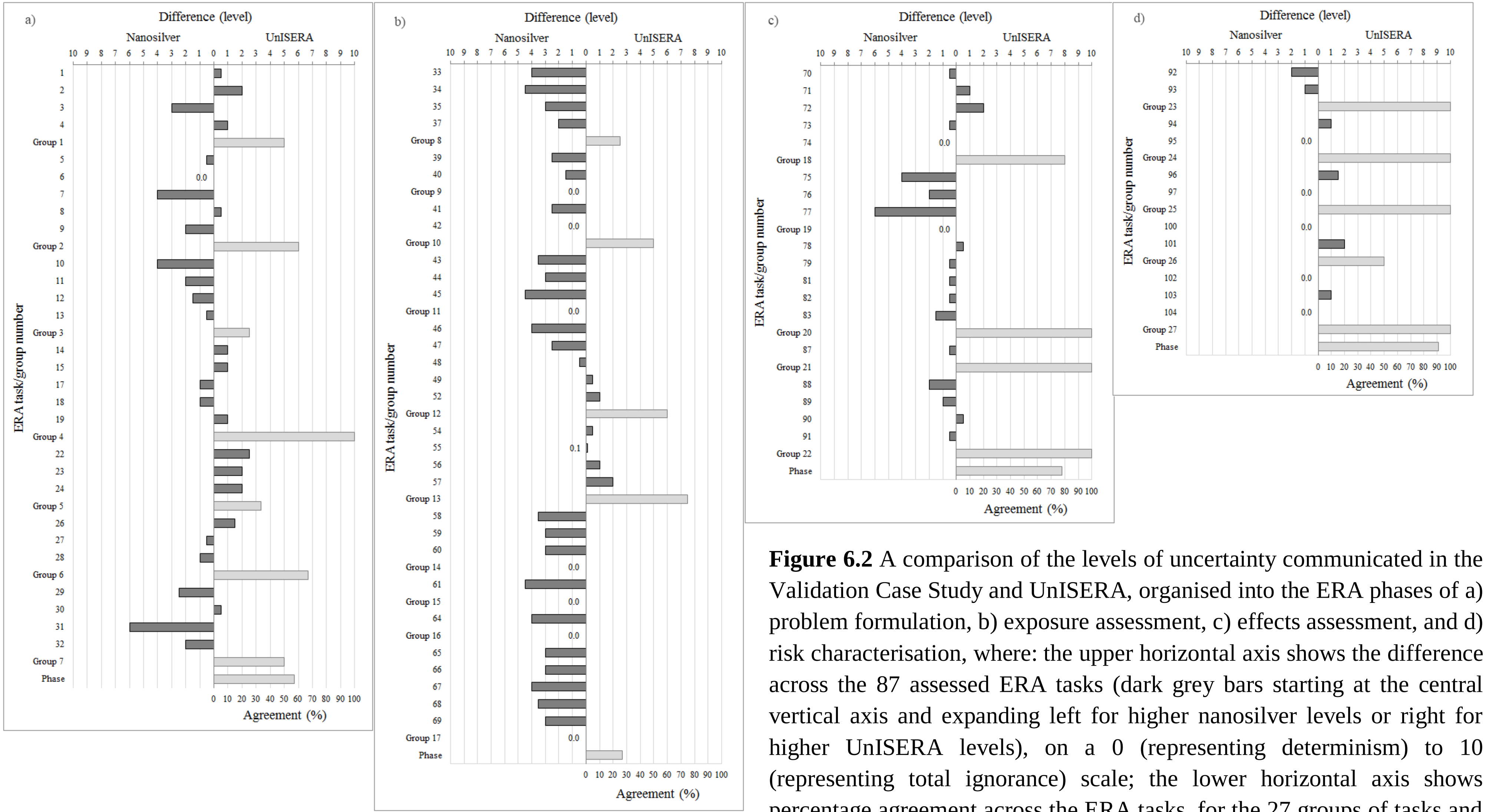


Figure 6.2 A comparison of the levels of uncertainty communicated in the Validation Case Study and UnISERA, organised into the ERA phases of a) problem formulation, b) exposure assessment, c) effects assessment, and d) risk characterisation, where: the upper horizontal axis shows the difference across the 87 assessed ERA tasks (dark grey bars starting at the central vertical axis and expanding left for higher nanosilver levels or right for higher UnISERA levels), on a 0 (representing determinism) to 10 (representing total ignorance) scale; the lower horizontal axis shows percentage agreement across the ERA tasks, for the 27 groups of tasks and overall for the four ERA phases (light grey bars). The ERA tasks are separated into the groups listed in Table 6.3.

The nature dimension has three aspects to consider. The epistemic category agreed in 26 out of 28 tasks, the aleatory category in 28, and the combined category in 22, for a median agreement across this dimension of 26 out of 28 tasks (93%). When considering the difference in occurrence rates between the Validation Case Study and UnISERA (Figure 6.3), the epistemic category returned a median difference of 16%, the aleatory category of 10%, and the combined category of 23%, for a median difference in occurrence rates across all tasks in problem formulation of 14%.

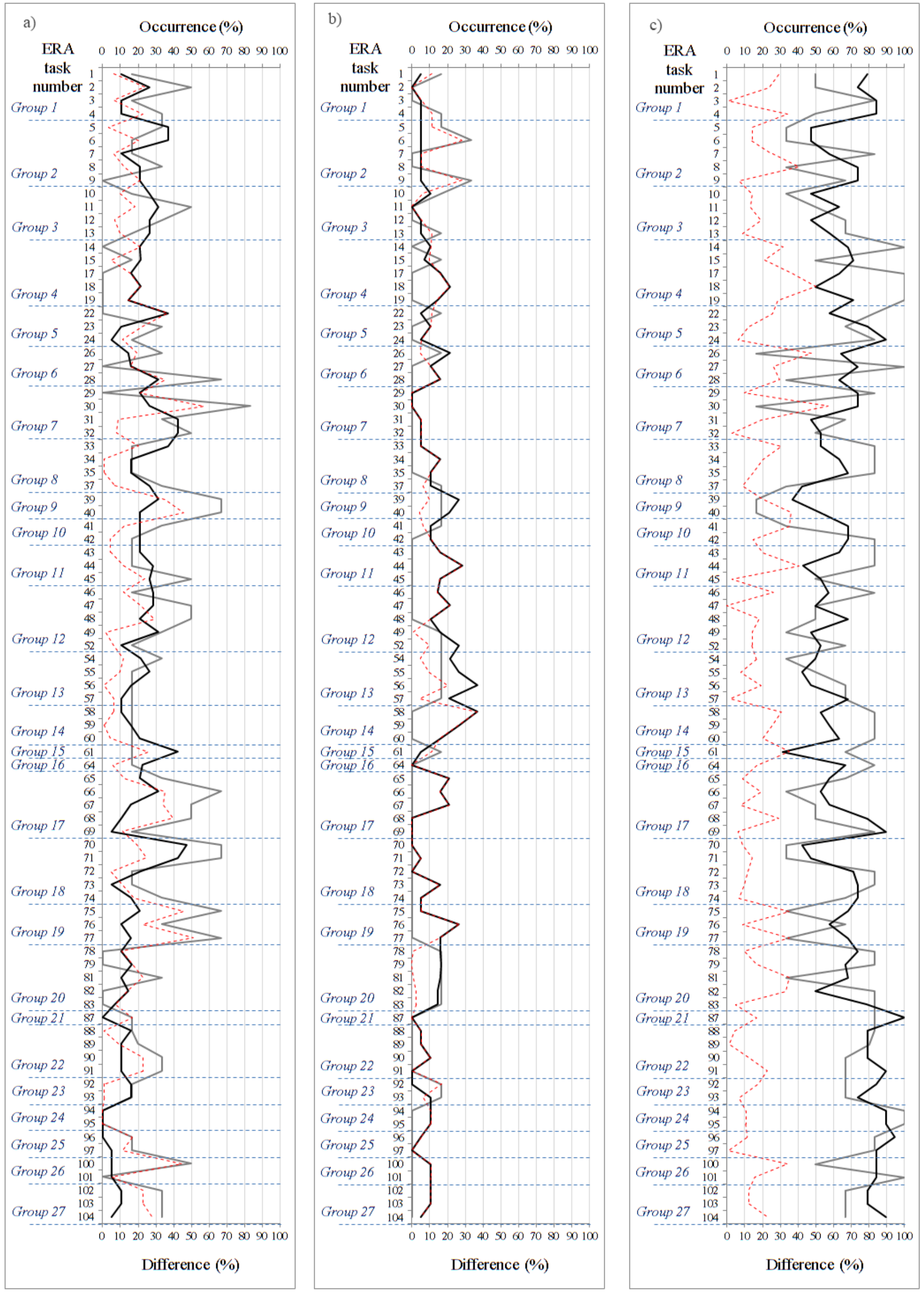


Figure 6.3 A comparison of occurrence proportions (%) between the Validation Case Study (solid grey lines) and UnISERA (solid black lines) across the 87 validated ERA tasks, for the nature-based uncertainties of a) epistemic, b) aleatory, and c) combined. Difference (%; dashed red lines) is also shown on the lower horizontal axis. The ERA tasks are separated into the groups listed in Table 6.3.

The median agreement across all seven locations was 20 out of 28 tasks (71%). The language category showed the highest agreement (27 out of 28 tasks), and model uncertainty the lowest (18 out of 28). The experts particularly disagreed on the extent to which decision uncertainty was present in group 7, identifying data considerations, with a median difference of 37% (Figure 6.4). The median difference in occurrence rates across all tasks and all seven locations was 20%.

Exposure assessment

The median level of uncertainty in the Validation Case Study was 6.0, compared with 4.0 for UnISERA. Exposure assessment yielded the lowest level of validation on a phase-by-phase basis (Figure 6.2b), with just 8 out of 30 tasks agreeing (27%). Of the seven groups across the four phases that were in complete disagreement (i.e. all tasks within the group disagreed), six were in exposure assessment, namely collecting information about the stressor's composition (group 9; P=0.01) and release (group 11; P=0.00), evaluating stressor-receptor co-occurrence (group 14; P=0.00) and contact (group 15; P=0.02), integrating multiple lines of evidence (group 16; P=0.02), and creating exposure profiles (group 17; P=0.00). However, group 13, collecting information about the receptor, showed agreement across its tasks (75%; P=0.01).

The median agreement between the Validation Case Study and UnISERA across the three categories of the nature dimension was 26 tasks out of 30 (87%). Whilst the level of agreement was higher in the previous phase, the difference in relative occurrence percentages across this dimension was actually 2% lower here, at 12% (Figure 6.3). A median occurrence difference of 0% was even observed for the aleatory category in group 16, which concerned the integration of multiple lines of evidence.

Exposure assessment provided UnISERA with the highest level of validation across the four phases for the location dimension, with 27 out of 30 tasks in agreement (90%). Language uncertainty agreed in all 30 tasks, whilst data and model uncertainties each agreed in just 21. These agreement levels are also reflected in the median differences in group-level occurrence rates for the different locations, with language uncertainty showing a difference of 0% in groups 13, 14 and 15, and data uncertainty differing to an extent of 58% in group 15. The median difference in occurrence rates across all tasks and locations in exposure assessment was 17%, the lowest of the four phases (Figure 6.4).

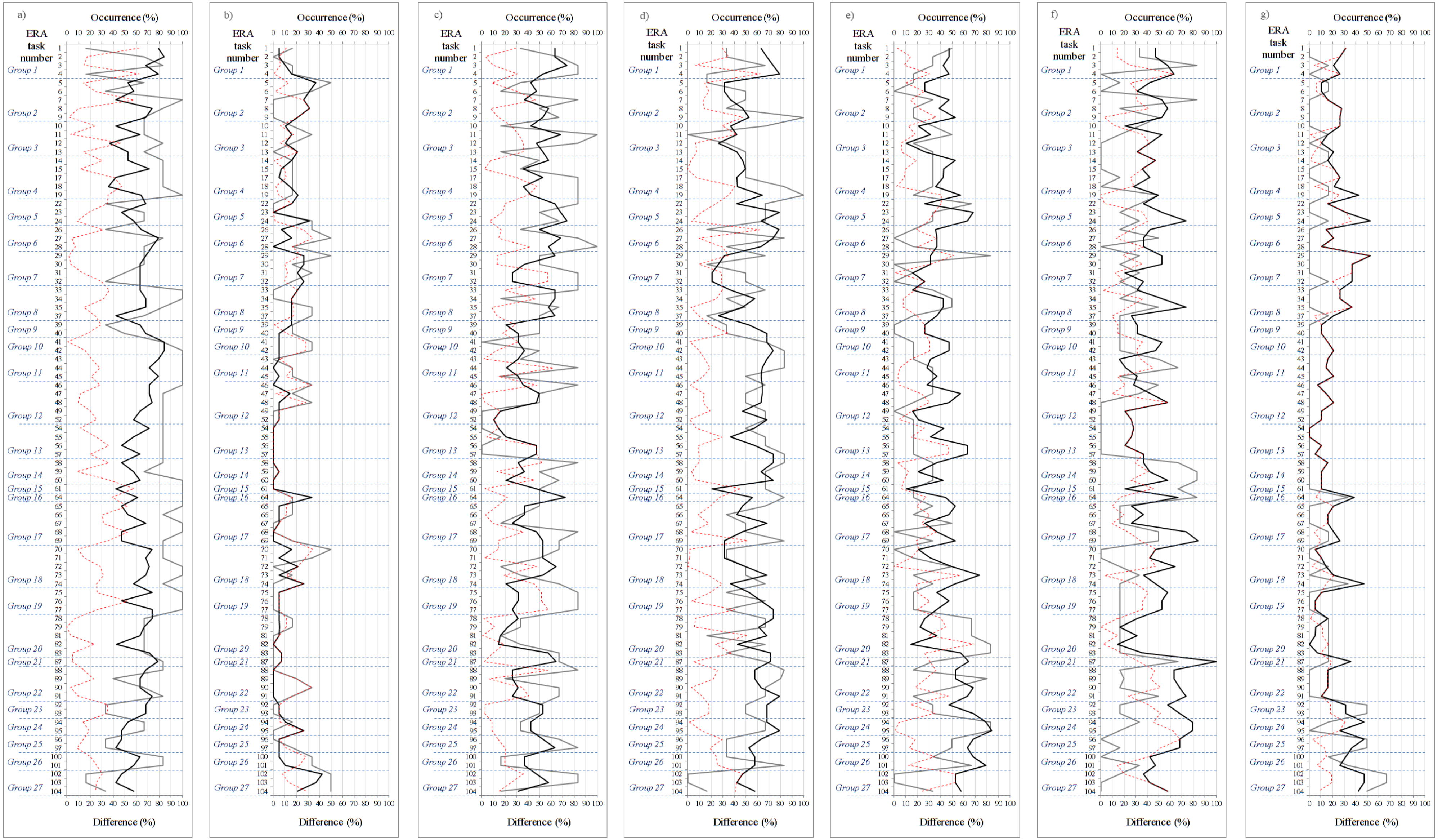


Figure 6.4 A comparison of occurrence proportions (%) between the Validation Case Study (solid grey lines) and UnISERA (solid black lines) across the 87 validated ERA tasks, for the location-based uncertainties of a) data, b) language, c) system, d) variability, e) extrapolation, f) model, and g) decision. Difference (%; dashed red lines) is also shown on the lower horizontal axis. The ERA tasks are separated into the groups listed in Table 6.3.

Effects assessment

The median level of uncertainty in the Validation Case Study was 5.5, compared with 4.3 for UnISERA. A higher degree of agreement was seen for the level dimension in effects assessment than in the previous phases (Figure 6.2c), with 14 out of 18 tasks (78%) falling within the validation range. There were three separate groups of tasks that returned 100% agreement rates: group 20, assessing the effect endpoints (*P*=0.88); group 21, integrating multiple lines of evidence (*P*=0.77); and group 22, creating the stressor-response profile(s) using single point and distribution-based methods (P=0.81). Conversely, just one group showed a 100% level of disagreement (group 19; *P*=0.00).

A median of 16 out of 18 tasks (89%) were in agreement across the three categories of the nature dimension in this phase. The aleatory category showed a median agreement of 100%, and a median difference in occurrence rates of just 4% across the 18 tasks (Figure 6.3). The overall difference in occurrence rates across the dimension in this phase was the same as for exposure assessment, at 12%.

Exposure assessment saw a median agreement in 15 out of 18 tasks (83%) across the seven locations of uncertainty. Data and language uncertainties agreed to the highest extent (both 17 out of 18 tasks), with model uncertainty again showing the lowest level of agreement (9 out of 18 tasks). Model uncertainty also contained the largest difference in occurrence rates across this phase, at a median of 35% (Figure 6.4). Conversely, data uncertainty tallied more closely across the validation study and UnISERA, with a median difference of 15%, and just 5% in groups 20 and 21.

Risk characterisation

The median level of uncertainty was 5.0, the same as for this phase in UnISERA. Risk characterisation recorded the highest level of validation of the four phases (Figure 6.2d), with 10 out of 11 tasks agreeing (91%). Four out of the five groups in this phase had agreement levels of 100%: group 23, selecting relevant exposure and effects profiles (*P*=0.75); group 24, estimating risk levels (*P*=0.16); group 25, aggregating risk levels (*P*=0.12); and group 27, assessing the significance of the risk levels (*P*=0.32). Five of the 11 tasks in this phase also shared the same median values in the Validation Case Study and UnISERA, namely tasks 95 (*P*=0.40), 97 (*P*=0.59), 100 (*P*=0.80), 102 (*P*=0.44) and 104 (*P*=0.95).

Risk characterisation showed a median agreement for the nature dimension of 10 out of 11 of its tasks (91%). The difference in occurrence rates was lower in this phase than any other, at 11%, with median differences of 11% for the epistemic category, 11% for aleatory, and 12% for combined (Figure 6.3).

Despite large variation in the values across the seven locations, risk characterisation yielded a median agreement in 9 out of 11 tasks (82%). Language and decision uncertainty both returned an agreement of 100%, with model uncertainty agreeing in just 18% of cases (2 out of 11 tasks). Whilst model uncertainty also contained a high median difference in occurrence rates of 42%, the value for extrapolation uncertainty was even higher, at 53% (Figure 6.4). No other location, on an ERA-phase basis, differed more between the Validation Case Study and UnISERA. However, the largest difference on a group-basis, also seen in this phase, was for model uncertainty, at a median of 63%. It follows then that across all seven locations and 11 tasks, risk characterisation produced the largest difference in occurrence rates of any phase, with a median value of 20%.

Overall

The median level of uncertainty across all 99 tasks in the Validation Case Study was 5.0, compared to 4.0 for UnISERA. Both levels were therefore in the same range of scenario uncertainty. The highest degree of agreement for the level dimension was seen in risk characterisation and the lowest in exposure assessment, with an overall agreement (across all tasks in the four phases) of 55%. Across all 87 tasks, the median difference in the level of uncertainty between comparable values in the Validation Case Study and UnISERA was just 0.5.

Regarding the nature of uncertainty, the epistemic category had an overall agreement rate of 89% (with a median difference of 11% in occurrence rates), the aleatory category of 99% (with a median 11% difference), and the combined category of 84% (with a median 12% difference). Across all 87 tasks and all three nature categories the nature dimension had a median agreement of 90%, and a median difference in occurrence rates of 12%.

The location dimension, across its seven categories, yielded a median agreement of 80% over the four phases. The highest degree of validation was seen for the locations of language (98%) and decision (94%), and the lowest for model (58%) and system (71%). Over the four

ERA phases, the difference between the values in the Validation Case Study and the comparable values in UnISERA was highest for model uncertainty, with a median difference of 45%, with a median difference of 18% across all seven locations of uncertainty.

6.4 Discussion

6.4.1 Uncertainty across the phases of the Validation Case Study

Problem formulation

The problem formulation phase of the Validation Case Study contained the lowest levels of uncertainty across the four phases, with a median value of 4.0, and generally low occurrence rates for the categories of the nature and location dimensions. However, group 7, which considered the data requirements for the ERA, had a relatively high median level of 5.8, and strong associations with the epistemic data and system uncertainties. This corresponds with recent concerns raised about the availability, specificity, and reliability of the databases used in ENM ERAs (Mueller and Nowack 2008; Aschberger *et al.* 2011; Musee 2011). A lack of understanding about the measurements and data required has also been demonstrated (Hristozov *et al.* 2012). However, the task with the highest level of uncertainty in problem formulation (7.0; task 3) related to using existing evidence at the beginning of the phase to better constrain potential exposure scenarios.

Exposure assessment

Exposure assessment had the highest associated level of uncertainty of the phases in the Validation Case Study, with a median value of 6.0. Some of the most dominant levels, natures and locations of uncertainty concerned collecting information about the composition, distribution and release of the stressor (groups 9, 10 and 11). That uncertainty is associated with the chemical and physical composition of a nanomaterial is well known (Chen *et al.* 2011; van Broekhuizen *et al.* 2012). Here, whilst experts did not communicate that these aspects (group 9) contained the same high levels seen elsewhere (with a median value of 4.5), they did note their existence was primarily due to a lack of knowledge (with a median epistemic occurrence rate of 67%), and specifically related to sub-optimal system-based understanding.

Groups 10 and 11 also contained epistemic uncertainty through high levels of the data location. The tasks in these two groups aimed to determine the spatial and temporal distribution of the stressor and the probability, quantity, and intensity of its release, and are the required components of a PEC (van der Oost *et al.* 2003). Formulation of a PEC for nanomaterials is not straightforward, largely because particle aggregation can lead to a range of sizes and distributions, and ultimately a range of concentrations (Gottschalk *et al.* 2011). Suitable datasets that specifically describe ENM PECs in natural systems are, at present, lacking (Quik *et al.* 2011), or are often only available through manufacturers' reports at national scales (Gottschalk *et al.* 2010). The strong connection to data uncertainty was therefore expected.

Groups 10 and 11 also contained high rates of variability uncertainty, consistent with recent research which found that variability had a big effect on PECs of ENMs in a river environment (Gottschalk *et al.* 2011). It is also common practice to use models to formulate PECs across environmental media (Boxall *et al.* 2007, Mueller and Nowack 2008, Gottschalk *et al.* 2009), though the procedures are often based on assumptions and are susceptible to uncertainty in the model's output (Aschberger *et al.* 2011). It is therefore a little surprising that the modelling location-based uncertainty only returned a median occurrence rate of 17% across groups 10 and 11, though the experts may have instead apportioned their model-based concerns to the epistemic data and aleatory variability locations, since model uncertainty (which has a combined nature of epistemic and aleatory) can contain both of these types of uncertainty (see Section 4.4.1).

The fate and behaviour of ENMs in different environmental compartments is another potentially uncertain area of exposure assessment that requires consideration (Chen *et al.* 2011; Musee 2011). Here, the tasks associated with the biological, chemical, and physical aspects of a stressor's fate and transport (numbers 46, 47 and 48) had reasonably high median occurrence rates of data (83%) and system (50%) uncertainties. Whilst a full understanding of ENM fate in natural systems, including freshwater environments, does not yet exist (Gottschalk *et al.* 2011), steadily expanding insight is driving further research in this area (Eckelman *et al.* 2012), which may in time help to reduce the associated epistemic uncertainties.

The group associated with determining the spatial, temporal and intensity of overlap between the stressor and receptor (group 14) had the joint-second highest level of uncertainty of the groups in exposure assessment, with a median value of 7.5, and was heavily associated with the combined nature category, manifesting in data, system, variability and model uncertainties. Furthermore, across all seven locations, group 14 returned a median occurrence rate of 67%, the highest of any group in the case study. The high levels of uncertainty seen in all three dimensions here is largely attributed to the fact that relatively little is known about the release of ENMs into the environment, and the effects that different media have on their behaviour (Gottschalk and Nowack 2011). Although firmly part of the exposure assessment phase, when assessed within a controlled environment, such as a laboratory, group 14 shares similarities with the set of tasks in effects assessment associated with determining the duration, frequency and intensity with which the receptor was exposed to the stressor (group 19).

Effects assessment

The tasks in group 19 contained the highest median level of uncertainty (9.0) of any of the 27 groups across the Validation Case Study. The experts concluded that the associated uncertainty was predominantly epistemic in nature, and manifest in the locations of data and system. The three tasks in group 19 can be difficult to determine with any great precision, especially for mobile species such as fish, and often require the use of best-estimates (van der Oost *et al.* 2003), allowing epistemic uncertainty to exist. Aleatory uncertainty impacted group 19 far less than it did group 14, highlighting the differing extent to which aleatory processes contribute to measurements in controlled and natural environments.

Risk characterisation

Whilst the first three phases of an ERA largely call for a defined protocol to be followed, the risk characterisation phase encourages more independent decisions to be made. For example, the inclusion of relevant exposure and effects profiles (group 23), the aggregation of certain risk estimates (group 25), and the selection of appropriate thresholds against which to judge the significance of the risk (group 27), are all dependent upon the choices of the risk analyst(s). Consequently, the risk characterisation phase contained far higher median occurrence rates of decision uncertainty (50%), than problem formulation (0%), exposure assessment (0%), or effects assessment (8%).

Overall

Across all 99 tasks, the median level of uncertainty in the Validation Case Study was 5.0, firmly in the range of scenario uncertainty. At this level outcomes of events, such as ERA tasks, can generally be defined, but the likelihoods associated with their occurrence remain largely unknown (see Section 2.7). In the context of ERAs, and the established equation of Risk = Likelihood X Consequence, uncertainty in the likelihood of an event occurring translates to higher levels of uncertainty in the tasks within the exposure assessment phase than elsewhere, as was the case here. Furthermore, unless non-epistemic locations were of major concern, an improvement in system knowledge coupled with a reduction in data uncertainty, which was the dominant location across this case study, would help to drive the high levels seen in exposure assessment, and elsewhere, farther from ignorance and closer to a position of determinism. This is the goal of uncertainty management, and a principal rationale behind UnISERA.

6.4.2 The appropriateness of the Validation Case Study

The purpose of validation is to test a system, method, or set of results against potential uses, whilst ensuring that the validation tests are applicable and fit for purpose (González and Herrador 2007). Ideally, a range of validation scenarios would be used, but the limited timeframe within which this research was conducted meant than only a single Validation Case Study was possible. It was therefore vital that the output from UnISERA was tested against a subject domain that carried the most validation-based benefits.

Although there exists discussion in the literature about the way in which engineered nanomaterials are best assessed (Rocks *et al.* 2008; Aschberger *et al.* 2011), the controlling legislation, such as REACH in the EU and the Toxic Substances Control Act (TSCA) in USA, recommends that accepted ERA methods are applied. In terms of the risk-based approach, the chosen validation case was therefore aligned with those in UnISERA, making it applicable in this context.

Validating the information within UnISERA against a risk domain directly comparable to those that comprised UnISERA would have served to validate the method, results, and observations against similar scenarios. This kind of approach is useful when an extensive validation program is permitted. However, in the context of a single-study validation, it can be more useful to utilise a test case that is similar in structure, but different in other ways (González and Herrador 2007). This approach recognises the importance of validating the output from UnISERA for application to a new domain, rather than re-application to the same domain(s), and therefore serves to make the validation process as realistic and useful as possible. A distinction on the basis of the quantity of empirical evidence available (i.e. established risk validated against emerging risk) was therefore deemed appropriate.

The selection of the four case studies (and specific risk-relationships) has the potential to alter the output from UnISERA (see Section 5.8.4) and the success of its subsequent validation. However, this would be true of any case studies. By establishing, following, and justifying the outlined selection criteria, this research has sought to be transparent and reproducible.

6.4.3 Validating the observations from UnISERA

The level of uncertainty

Of the three dimensions, the level of uncertainty had the lowest rate of agreement across all 87 tasks, at 55%. The Validation Case Study returned higher levels of uncertainty for 53 tasks, UnISERA for 26 tasks, with 8 tasks of equal value. Whilst this illustrates that the levels of uncertainty were higher in the validation study, it does not necessarily disprove the validity of the output from UnISERA, provided that this disparity can be explained.

The three case studies that comprise UnISERA were selected in part because of the quantity of information (i.e. ERAs) associated with them. It was important to build the system on highly researched subject domains, so that each expert could consider the (potentially non-intuitive) uncertainty-based aspects with a validated ERA scenario in mind. However, one of the key aims of UnISERA is to provide guidance in areas where the available quantity of information may be low. Whilst the choice of the validation subject domain was therefore relevant (see Section 6.4.2), some disparity in results was expected due to the currently limited extent to which the emerging risk domain has been researched. For example, the ENM exposure assessment phase, which shared an agreement rate of just 27% with the comparable phase in UnISERA, has relatively small quantities of information associated with aspects such as PEC determination (Quik *et al.* 2011), ENM fate and behaviour (Gottschalk *et al.* 2011), and stressor-receptor co-occurrence (Gottschalk and Nowack 2011). Experts participating in the Validation Case Study offered similar insights, including:

"at low level concentrations we just don't know if we have silver nanoparticles or other forms of silver. Thus, I think that we know nearly nothing about actual exposures";

and that the high level of uncertainty associated with determining stressor-receptor co-occurrence

"is largely attributable to the fact that we know so little about the release and movement of silver nanoparticles in the environment".

Therein lies the obvious disadvantage of using an emerging risk domain as a validation case: the stressor is novel, and its characteristics, release, and actions on environmental compartments as well as potential receptors are largely unknown. It follows that the lowest levels of agreement were associated with aspects involving the stressor, and the highest levels of agreement were seen for those aspects in which the stressor did not feature. For example, group 13, which sought to collect information about the receptor, returned the highest agreement rate across the exposure assessment phase (75%), whilst group 11, collecting information about the stressor's release, and group 14, determining stressor-receptor co-occurrence, yielded rates of 0%.

High rates of agreement were not only confined to aspects involving the receptor. Risk characterisation, for example, which draws together the output from the exposure and effects assessment phases, saw an overall agreement of 91% across its contained tasks. The Validation Case Study also matched UnISERA in terms of its median level of uncertainty in this phase, at 5.0. This observation that uncertainty levels can differ so much between different parts of the same assessment (e.g. between exposure assessment and risk characterisation) supports the view that uncertainties should be first dealt with in the phase in which they occur (Janssen *et al.* 2003; Refsgaard *et al.* 2007), rather than leaving uncertainty analysis as a task for risk characterisation (US EPA 1998; Fairman *et al.* 1998; Defra 2011).

The outlined agreement rates were largely corroborated by the statistical measures applied to the data associated with each task. For example, of the 20 tasks with the lowest statistical significance, 14 were from exposure assessment. As for the 20 tasks with the highest significance values, nine were from effects assessment and six were from problem formulation, which had overall agreement rates of 78% and 57%, respectively. When using statistical tests that compare measures of central tendency, such as Mann-Whitney, it is important to note that they test for disagreement rather than agreement. Therefore, a low *P*-

value (towards 0.00) signifies disagreement across the tested groups, and a high *P*-value (towards 1.00) signifies agreement. Three tasks with a *P*-value of 0.00 were recorded: task 31, concerned with identifying the required data collection techniques for the assessment; task 46, collecting necessary data for the biological aspects of fate and transport; and task 77, determining the intensity (i.e. concentration) of the stressor to which the receptor is exposed. These three tasks, all of which were firmly related to the stressor, had far higher levels of uncertainty in the Validation Case Study than in UnISERA. One task with a *P*-value of 1.00 was also recorded, namely task 74, which focussed on prioritising either empirical or experimental data during effects assessment, something which both sets of experts associated with low levels of uncertainty.

All four ERA phases within both UnISERA and the Validation Case Study contained median values that were in the range of scenario uncertainty. Therefore, on a phase-by-phase basis, there was 100% agreement. However, uncertainty is neither identified nor managed at the phase level, so this observation is made more out of interest than because of its practical use.

The nature of uncertainty

Of the three dimensions, the nature of uncertainty had the highest rate of agreement across all 87 tasks, at 90%. There were few instances of disagreement, either within the three categories in this dimension, or overall. Certainly, there were no areas of sustained disagreement, with only intermittent tasks falling outside the defined 33% agreement range. The highest rate of agreement across all 87 tasks was observed for the aleatory category, with 99%. There was therefore only one task in which the difference in occurrence rates was more than 33%, namely task 58 (in exposure assessment), which sought to evaluate the spatial overlap between the stressor and the receptor. Here, a median value of 37% was returned for the aleatory category in UnISERA, with 0% recorded for the corresponding task in the Validation Case Study.

The generally small differences between the median occurrence rates of corresponding values in UnISERA and the Validation Case Study hold important meaning, but must also be framed with a caveat: similar values were often seen where occurrence rates are low. These categories were mutually exclusive; they could be selected on an individual basis, but not in tandem. For example, when respondents selected both the epistemic and aleatory categories the combined category was automatically invoked, whether or not either of its two locations (model and decision) were then selected. This ensured that, in such instances, it was recognised that both epistemic and aleatory uncertainty occurred simultaneously in the task being assessed, and that this important information was not diluted during the aggregation process. However, it did result in low values for the epistemic and aleatory categories in situations where they were both deemed to exist, and therefore high values for the combined category. Ultimately, this may have resulted in slightly inflated agreement rates, and slightly lower differences between occurrence rates, than may have been observed had the nature categories not been treated as mutually exclusive.

The location of uncertainty

Across its seven categories, and all 87 validated tasks, the location dimension returned a median agreement rate of 80%. The location that recorded the largest overall disagreement was model uncertainty, with a median value of 58%, resulting from the fact that the occurrence rates in UnISERA were far higher (with median values of 45%, 32%, 50% and 58% across the four phases) than the rates seen in the Validation Case Study (with median values of 17% in all four phases). The experts who participated in the Validation Case Study therefore did not believe that model uncertainty was a big concern. Instead, they returned higher overall rates for the data and system locations than were found in UnISERA. This is linked to the debate about the quantity and quality of data and system-based understanding in ENM ERAs (Gottschalk and Nowack 2011). Strongly aligned to this debate, system uncertainty was the second least-validated location behind model uncertainty, with an agreement rate of 71%. However, the opposite relationship existed to that of the model location, since system uncertainty returned much higher values in the Validation Case Study than were seen in UnISERA, another effect of the disparity between the respective sizes of the available evidence bases.

The highest levels of agreement were seen for the language and decision locations, with median values of 98% and 94% across all 87 tasks. As with the nature dimension, extremely high levels of agreement were noted here because of the similarly low frequencies with which the respective uncertainties occurred. However, unlike the nature dimension, these low values were not the result of any mutual exclusivity between the categories.

A qualitative validation of the location dimension can also be performed by comparing the expert values in UnISERA to the frequencies of location-based uncertainties recorded during the analysis of 171 WOE ERAs in Chapter 4 (see Figure 4.2; Table 6.5).

Table 6.5 Occurrence percentages (and rank), for the uncertainties within the location dimension of UnISERA and the WOE ERA evidence base in Chapter 4 (see Figure 4.2), with % difference (and rank difference).

Location	UnISERA occurrence % (rank)	Chapter 4 occurrence % (rank)	Difference % (rank difference)
Data	22.1 (1)	32.5 (1)	10.4 (=)
Language	4.0 (7)	4.2 (6)	0.2 (1)
System	15.4 (4)	13.0 (3)	2.4 (1)
Variability	19.6 (2)	12.2 (4)	7.4 (2)
Extrapolation	15.1 (5)	28.4 (2)	13.3 (3)
Model	16.2 (3)	8.1 (5)	8.1 (2)
Decision	7.6 (6)	1.6 (7)	6.0 (1)
TOTAL	**100.0**	**100.0**	

According to this qualitative comparison, the most highly validated location was language uncertainty, with a difference of just 0.2% between its occurrence proportion in UnISERA and in the WOE ERA evidence base. The associated values were both low, similar to the validation performed in this chapter, allowing for near parity. Extrapolation uncertainty returned the lowest degree of validation, with a difference of 13.3%, and a difference in rank order position of three places. Both extrapolation and data uncertainties returned far lower proportions in UnISERA than in the WOE ERA evidence base, reflective of the fact that they were the two biggest location-based concerns in the articles reviewed. However, an exploration into the factors affecting these proportions in the WOE ERA evidence base is not possible, since the 171 articles were drawn from a number of disparate research domains.

This qualitative comparison and quantitative validation suggest that language and decision uncertainties can be highly validated on a consistent basis whilst the other locations may vary more, an effect of their higher occurrence rates. The degree to which the different aspects of the three dimensions are validated is likely to be affected by the metrics used in the process, and the range of acceptability adopted.

6.4.4 Modifying the validation agreement range

Two distinct validation metrics were used in each uncertainty dimension; one that was user-defined and one that was more rigid. The rigid techniques, termed so because they are standard validation metrics for non-uniform datasets, were the Mann-Whitney statistical test for comparison of level-based central tendencies (i.e. medians), and comparison of median occurrence rates for the nature and location dimensions. Whilst these rigid techniques are appropriate, they are utilised here as secondary measures of validation, with agreement ratios and percentages the more important gauge. Here, agreement occurred on the basis of user-defined maximum thresholds. The purpose of this sub-section is to explore how variation in the user-defined metrics affect the validation outcome.

The user-defined metric for the level dimension assessed agreement on the basis of like values (in UnISERA and the Validation Case Study) belonging to the same established categories of statistical, scenario or ignorance-based uncertainty. Agreement was also noted where there was a maximum difference of 1.0 (i.e. 10% of the 0 to 10 level scale) between corresponding values, to avoid a situation where two data-points could be similar to each other yet in different level categories and therefore in disagreement. Removing this extra 1.0 tolerance and defining agreement on the basis of the three categories alone has an effect on the validation results. Under this new validation scenario, problem formulation sees a shift in agreement across all its tasks from 57% to 36%, exposure assessment from 27% to 23%, effects assessment from 78% to 61%, and risk characterisation remains at 91%. Across all 87 validated tasks, the agreement alters from 55% to 44%. The categories of statistical, scenario and ignorance-based uncertainty were introduced (along with determinism and total ignorance) by Walker *et al.* (2003) and applied to an ordinal scale of zero to one by Krayer von Krauss *et al.* (2004). An argument could be made for placing another category along this scale, one where the probabilities of an event occurring can be constrained, but the outcomes remain unknown; such a category is discussed elsewhere, and referred to as 'ambiguity' (Stirling *et al.* 1999; see Section 2.7.3). Its inclusion would shift the existing numerical boundaries between the groups by 8% (e.g. from 33% to 25%), similar to the additional 10% applied in this research. Therefore, there is merit in providing this extra tolerance when assessing agreement of level-based values.

The user-defined metric for the nature and location dimensions was similar to that of the level, in that a category size of 33% was established. Whereas the level dimension called for three fixed categories (statistical, scenario, ignorance) each of 33%, agreement in the nature and location dimensions was assessed using a mobile bracket of ±33% around the value being validated. This threshold was chosen because it represented the maximum percentage extent to which two values in the level dimension could differ (e.g. 1% and 33%) whilst still being in agreement. Employing this range for nature- and location-based values aids consistency across the dimensions, but also offers some scope for variation. Altering the agreement range of nature-based values to ±25%, for example, would make little impact, with agreement across all 87 tasks changing from the existing rate of 90% to 86%. However, a shift to ±20% would see the overall agreement drop considerably to 70%. The effects in the location dimension would be more pronounced, with a decrease from the existing agreement value of 80% down to 64% when the ±25% bounds is applied, and to 56% in the case of ±20%.

Whilst interesting to note the effects of such modified ranges of acceptability, there is no basis for applying anything other than a value of ±33%. As such, the validation exercise performed in this chapter is deemed full and appropriate.

6.5 Conclusion

The purpose of this chapter was to validate the output from UnISERA. The domain of engineered nanomaterials was selected, according to outlined criteria, along with a specific validation scenario of consumer-based nano-Ag risk to freshwater fish. A set of validation results was created for this scenario using the output from structured elicitations, in which six subject-matter experts assessed ERA templates for the three dimensions of uncertainty: level, nature, and location. The results from the validation scenario were compared to the corresponding values within UnISERA's output. Validation was judged primarily on the basis of corresponding values falling within pre-determined ranges of acceptability, yielding agreement percentages across the different tasks, groups of tasks, and phases within UnISERA. Statistical analyses (Mann-Whitney) and comparisons of median occurrence rates were also performed. Of the 89 tasks in UnISERA, 87 were also present in the validation scenario, with 28 in problem formulation, 30 in exposure assessment, 18 in effects assessment, and 11 in risk characterisation. Validation was therefore performed on these 87 tasks.

The highest degree of agreement for the level dimension was seen in risk characterisation (91%) and the lowest in exposure assessment (27%), with an overall agreement across all 87 tasks of 55%, the lowest rate seen across the three dimensions. However, the median difference in the level of uncertainty between comparable values in the Validation Case Study and UnISERA was just 0.5, highlighting a closer parity than the agreement rate may suggest.

Regarding the nature of uncertainty, the epistemic category had an overall agreement rate of 89%, the aleatory category of 99%, and the combined category of 84%. Across all 87 tasks and all three categories the nature dimension had a median agreement of 90%, the highest of the three dimensions, the highest of which was seen in problem formulation (93%), and the lowest in exposure assessment (87%). The location dimension, across its seven categories, yielded a median agreement of 80% over the four phases, with the highest in exposure assessment (90%), and the lowest in problem formulation (71%). For the individual locations of uncertainty, the highest degree of agreement was seen for language (98%) and decision (94%), and the lowest for model (58%) and system (71%).

The research in this chapter has drawn attention to the aspects of UnISERA's output that agreed and disagreed with the results from the chosen validation scenario. Both the nature and location dimensions within UnISERA showed high degrees of validation, with the exception of some aspects of the model and system locations. However, the level dimension was validated to a lesser extent, especially for tasks relating to the stressor material. These observations will help to drive future research, which is required in order to expand, re-validate and test UnISERA's method and findings, and will also offer insights into the appropriateness of subsequently selected uncertainty management techniques.

[This page is intentionally left blank]

Chapter 7: Summary and conclusions

7.1 Research aim and objectives restated

The aim of this research was to understand why ERAs can fail to identify uncertainties, and to provide a novel approach for addressing the identified issues. The main objectives of this research were to:

i. evaluate critically the primary methods for characterising and identifying uncertainty (i.e. typologies) in environmental risk-based systems, also considering their application to ERAs;
ii. create an evidence-based typology that drew from the existing set of peer-reviewed ERAs and addressed the issues raised in objective i), whilst also investigating the connections between uncertainties and other aspects of the ERA process;
iii. elicit experts' views on the types and magnitudes of uncertainty present within empirical (i.e. evidence-heavy) risk domains (applying the developed typology from objective ii), leading to the creation of a generic uncertainty identification system that was organised by the different stages and tasks within an ERA; and
iv. validate the generic uncertainty identification system against the elicited views of experts in an emerging (i.e. evidence-light) risk domain, highlighting areas of strength and weakness.

7.2 Thesis summary

The research in Chapter 3, which provided a review and analysis of 30 uncertainty typologies, aimed to achieve thesis objective (i). Several flaws were identified with the existing set of typologies used across environmental risk domains, namely that they: use terminology that is often contradictory; communicate varying frequencies and dimensions of uncertainties; source information from limited data sets; and are not appropriate for use, on an individual basis, with ERAs in order to characterise the range of potential uncertainties. The research in Chapter 3 identified the need to exercise caution when using existing uncertainty typologies due to their outlined limitations, and established the need for a new characterisation of potential uncertainties in ERAs.

The research in Chapter 4 aimed to address the limitations of the existing set of typologies, as indicated by thesis objective (ii). The research presented a typology of uncertainties based on the analysis of 171 peer-reviewed environmental WOE assessments, along with frequency and statistical analyses of the relationships between identified uncertainties and uncertainty management techniques as well as other aspects of the assessment process. This research produced the following main observations:

- seven locations of uncertainty exist within the examined evidence base (data, language, system, extrapolation, variability, model, and decision), with 20 related sub-types;
- some analysts are needlessly impacting the validity of their risk estimates by choosing not to manage uncertainty, though the majority of techniques used to manage the identified uncertainties are used appropriately;

The research in Chapter 4 emphasised the fact that whilst the new typology was appropriate for characterising uncertainties in ERAs, more could be done to aid the uncertainty identification process.

The research in Chapter 5 aimed to implement the typology introduced in the preceding chapter as part of a system designed to help identify uncertainties in ERAs, thus achieving thesis objective (iii). The developed uncertainty identification system for environmental risk assessments (called UnISERA) was based on the aggregated results of 19 structured expert elicitations across case studies encompassing the risk domains of genetically modified higher plants, particulate matter, and agricultural pesticides. The output from UnISERA described the uncertainty associated with 89 distinct tasks across the four phases of an ERA (28 in problem formulation, 32 in exposure assessment, 18 in effects assessment, and 11 in risk characterisation), allowing for the following main observations:

- risk characterisation contains the highest levels of uncertainty, where estimating, aggregating and evaluating risk levels are aspects of concern. Conversely, problem formulation contains the lowest levels of uncertainty, with the highest values related to choosing and considering the appropriateness of assessment endpoints;
- the combined epistemic and aleatory category is the dominant nature of uncertainty throughout the phases of an ERA;

- data is the dominant location of uncertainty in problem formulation, exposure assessment and effects assessment, and extrapolation is the dominant location in risk characterisation.

The research in Chapter 5 also acknowledges that the output from UnISERA can be reorganised to fit requirements, enabling analysts to prioritise ERA phases, tasks, and groups of tasks according to either the highest levels of uncertainty, the potential for the uncertainty to be reduced or only quantified, or the associated types of location-based uncertainty.

The research in Chapter 6 aimed to validate the statements made in the preceding chapter regarding the output from UnISERA, thereby achieving thesis objective (iv). The results from the validation scenario, which focussed on the domain of engineered nanomaterials and in which six subject-matter experts participated, were compared to the corresponding values within UnISERA's output. Validation was performed on 87 out of the 89 ERA tasks in UnISERA, resulting in the following main observations:

- the level of uncertainty has the lowest degree of validation of the three dimensions, with an overall agreement across all 87 tasks of 55%. The highest degree of agreement is seen in risk characterisation (91%) and the lowest in exposure assessment (27%);
- the nature of uncertainty has the highest degree of validation of the three dimensions, with an overall agreement across all 87 tasks of 90%. The epistemic category has an overall agreement rate of 89%, the aleatory category of 99%, and the combined category of 84%;
- the location dimension has a high degree of validation, despite containing seven categories, with an overall agreement across all 87 tasks of 80%. The highest degree of agreement is seen in exposure assessment (90%) and the lowest in problem formulation (71%). For the individual locations of uncertainty, the highest degree of agreement is seen for language (98%) and decision (94%), and the lowest for model (58%) and system (71%).

The research in Chapter 6 also underlines that acknowledging the highly-validated aspects of UnISERA provides analysts with valuable guidance relating to uncertainty identification in ERAs, and that acknowledging the poorly-validated aspects helps to drive future research, as discussed in this chapter.

Uncertainty analysis, in the context of ERAs, aims to identify potential uncertainties throughout the ERA process and implement tools for their management. The success of these tools (designed to quantify and/or reduce and/or remove uncertainties) can be hindered by uncertainties going unidentified, which in turn can result from an inaccurate characterisation of potential uncertainties. Therefore, uncertainty typologies used by risk analysts can directly affect the validity of the following uncertainty analysis. Furthermore, application of these typologies for uncertainty identification has serious flaws.

The different aspects of this research now provide environmental risk analysts with an understanding of the limitations of existing typologies of uncertainty (Chapter 3), a new evidence-based typology that can be applied to ERAs in order to characterise potential uncertainties (Chapter 4), and a prioritised set of potential uncertainties for consideration through the ERA process, drawn from a novel identification system (UnISERA). This research therefore aids the characterisation and identification components of uncertainty analysis, with respect to ERAs, and is significant for a number of reasons.

7.3 Research significance

7.3.1 Significance regarding the characterisation of uncertainties in ERAs

There are many examples of typologies of uncertainty relevant to environmental risk domains, as evidenced by the relatively large set analysed in Chapter 3. The research presented in that chapter highlighted the limitations of those typologies, particularly in the context of ERAs, and outlined the requirement for a new typology to be created in a different way to those in the existing set, which was the focus of Chapter 4. The novel uncertainty typology presented in this research is significant for several reasons.

Research suggests that no typology "includes all of its meanings in a way that is clear, simple, and adequate for each potential use of such a typology" (Petersen 2006). Furthermore, as the analysis in Chapter 3 described, none of the existing typologies was deemed applicable to ERAs. Significantly, analysts are now equipped with a single typology of uncertainties with which they can attempt to understand the potential types that can exist within ERAs. This single version will eliminate the need for ERA analysts to consult several other typologies, thereby reducing their exposure to contradictory terminology and varying frequencies and dimensions of uncertainty, which may previously have been a barrier to a proper understanding regarding uncertainty characterisation.

This typology is the first, in the context of environmental risk, to consult the large existing set of published ERAs, drawing characterisations of uncertainty from their peer-reviewed content. It therefore minimises researcher subjectivity, a noted flaw of previous typologies (Knol et al 2009a), allowing for a defensible characterisation.

7.3.2 Significance regarding the identification of uncertainties in ERAs

Prior to this research being conducted, the primary tool for identifying uncertainties in ERAs was the uncertainty typology (Morgan and Henrion 1990; van Asselt and Rotmans 2002; Knol *et al.* 2009a). The differing abilities and experience levels of ERA practitioners often resulted in these typologies being used inconsistently (Gillund *et al.* 2008; Knol *et al.* 2009a), which allowed uncertainties in ERAs to go unidentified (EEA 2007; Hart *et al.* 2007; Dale *et al.* 2008).

Some attempts had been made to turn the traditional typology format (i.e. a list of organised uncertainties and their related definitions) into a more useful tool for uncertainty identification. For example, Walker *et al.* (2003) introduced the concept of an uncertainty matrix, later reproduced by Janssen *et al.* (2003), that contained blank sections which the analyst was encouraged to complete using either qualitative or quantitative information relevant to their system under study. However, in this approach it was still the sole responsibility of the analyst to locate the defined uncertainties, since no specific system-related guidance was provided.

The matrix approach was extended to include some example aspects of systems in which the different dimensions of uncertainty may have been present (e.g. data uncertainty relating to measures of population exposure; Refsgaard *et al.* 2007; Knol *et al.* 2009a). However, these example aspects, which were drawn from the domains of environmental modelling and burden of disease assessments, respectively, were not representative of the full range of potential aspects that would require consideration in such systems. Furthermore, analysts conducting ERAs were not significantly aided by this guidance, due to the limited cross-over between the research domains, and were therefore in a similar position as when equipped with a typology or blank uncertainty matrix. A little more guidance was offered by Knol *et al.* (2009a) who suggested that their uncertainty typology should be used to:

1) identify sources of uncertainty, by either:
    a. analysing each step of the assessment and relating uncertainties from the typology to those steps; or
    b. considering the uncertainties in the typology and explaining where in the assessment these uncertainties may occur;
2) prioritise each uncertain element within the assessment according to their relative importance; and
3) select one or more suitable tools (i.e. UMTs) to further analyse the identified uncertainties.

Thus, should the analyst have been capable of using the typology or matrix to identify all relevant uncertainties within the system under study, they would still be required to prioritise those uncertainties – according, for example, to a chosen significance metric, or temporal and financial restrictions – before one or more UMTs could be implemented. Guidance related to the selection of one or more UMTs, as noted by Knol *et al.* (2009a), is sufficiently good. If an analyst requires guidance on managing identified uncertainties, it is reasonable to assume that they also require guidance on identifying those uncertainties in the first place. Up to now, this guidance has been lacking, and uncertainty identification has remained a weakness of the uncertainty management process.

The UnISERA method and results introduced in Chapter 5 and validated in Chapter 6 equip environmental risk analysts with detailed guidance regarding the locations, natures and levels of uncertainties that are associated with all aspects of the ERA process. Moreover, these aspects are based on validated system maps, a method that is gaining in popularity within the risk community (Krayer von Krauss *et al.* 2004; Gillund *et al.* 2008; Krayer von Krauss *et al.* 2008; Ravnum *et al.* 2012; Smita *et al.* 2012; Zimmer *et al.* 2012). Importantly, the observations resulting from UnISERA reduce the reliance on the analyst to identify and prioritise uncertainties before they can be managed.

7.3.3 Significance regarding the management of uncertainties in ERAs

Several uncertainty management studies exist (in an environmental risk context) that combine potential UMTs with different levels, natures, and locations of uncertainty (van der Sluijs *et al.* 2004; Refsgaard *et al.* 2007; WHO 2008; Knol *et al.* 2009a). This thesis presented a similar analysis, which extended to the nature and location dimensions of uncertainty (see Section 4.3.2). The outlined relationships, regarding uncertainties and UMTs, within the literature and this research can be combined to further enhance the guidance available to analysts when selecting one or more UMTs (Table 7.1).

Table 7.1 Appropriate uncertainty management techniques for use in conjunction with different combinations of uncertainty (after Sections 2.8 and 4.3.2 of this thesis, van de Sluijs *et al.* 2004, Refsgaard *et al.* 2007, WHO 2008, and Knol *et al.* 2009a).

Nature →	**Epistemic**			**Aleatory**		**Combined**	
Location → / **Level ↓**	**Data** (Availability; Precision; Reliability)	**Language** (Ambiguity; Underspecificity; Vagueness)	**System** (Cause; Process; Effect)	**Variability** (Natural; Human)	**Extrapolation** (Inter/Intra; Laboratory; Quantity; Spatial; Temporal)	**Model** (Structure; Output)	**Decision**
Statistical	CI; EE; LHS; MCS; PDF; SA;	EE; SI;	BBN; EE; SI;	EE; LHS; MCS; PDF;	EE; LHS; MCS; PDF;	BBN; Boot; EE; EP; LHS; MCS; PDF; SeA;	BBN; EE; MCDA;
Scenario	EE; FDC; FL; PBA; SA; ScA;	EE; FL; ScA; SI;	BBN; EE; FDC; ScA; SI;	EE; PBA; UF;	EE; PBA; UF;	BBN; CI; EE; EP; PBA; ScA;	AM; BBN; EE; MCDA; ScA;
Recognised ignorance	EE; FDC; FL; NUSAP; PBA;	EE; FL; SI;	EE; FDC; NUSAP; SI;	EE; PBA; UF;	EE; PBA; UF;	EE; NUSAP; PBA;	EE; PM;

With acronyms corresponding to the UMTs of:
AM - Adaptive management[1]; BBN - Bayesian Belief Network[1,5]; Boot - Bootstrapping[1]; CI - Confidence intervals[1,5]; EP - Error propagation[1,2,3]; EE - Expert elicitation[1,2,3]; FDC - Further data collection[1]; FL - Fuzzy logic[1,4]; LHS - Latin hypercube sampling[1,4]; MCS - Monte-Carlo simulation[1,2,3,4]; MCDA - Multi-criteria decision analysis[1]; NUSAP - Numeral, unit, spread, assessment, and pedigree[2,3]; PM - Precautionary management[1]; PBA - Probability bounds analysis[4]; PDF - Probability density function[1]; ScA - Scenario analysis[2,3,5]; SeA - Sensitivity analysis[1,2,3,4,5]; SI - Stakeholder involvement[2,3,5]; UF - Uncertainty factor[1].

Where superscript values denote the sources used to assign UMTs to different uncertainty combinations:
1: Section 2.8 and Section 4.3.2; 2: van de Sluijs *et al.* 2004; 3: Refsgaard *et al.* 2007; 4: WHO 2008; 5: Knol *et al.* 2009a.

The structure of Table 7.1 is comparable to the uncertainty matrix templates discussed previously (Janssen *et al.* 2003; van der Sluijs *et al.* 2004; Refsgaard *et al.* 2007), albeit with updated natures and locations of uncertainty (drawn from the novel typology in Table 4.2) and an expanded set of UMTs. Prior to this research, the appropriate application of UMTs relied on the ability of the risk analyst to identify successfully the uncertainties that required managing. As explained, the research presented in this thesis has reduced that requirement, and, by extension, has put analysts in a better position to select one or more appropriate UMTs: analysts will now know which uncertainties can be expected to exist, and where to expect them, throughout the ERA process.

The information in Table 7.1 can be combined with the output from UnISERA (see Section 5.7.2) to form a complete set of guidance that helps analysts characterise, identify, and prioritise uncertainties within ERAs, and further provides suggestions for the selection of relevant UMTs to quantify, reduce or remove those uncertainties. An example of this is provided in Table 7.2, which combines the 10 ERA tasks (out of 87 tasks in total) with the highest median levels of uncertainty within UnISERA (see Table 5.14), along with their associated natures and locations of uncertainty, with appropriate UMTs.

Table 7.2 The 10 ERA tasks with the highest median levels of uncertainty within UnISERA, with accompanying ranked occurrence rates (with median values of at least 50%) for the nature and locations of uncertainty, and relevant corresponding uncertainty management techniques. The dimensions of uncertainty are shaded green where validation was successful, and red where not (see Sections 6.3.4 and 6.4.3).

ERA Task #	ERA Phase	ERA sub-phase	ERA task group	ERA task	Level[a]	Nature[b]	Location(s)[c]	Uncertainty management techniques (UMTs)[d]
72	Effects	Use available evidence to better constrain…	18. (Use available evidence to better constrain…)	Secondary stressors	7.0 (Ig) *P*=0.48	Co	1: Dat 2a: Sys 2b: Mod 3a: Var 3b: Ext	1: EE, FDC, FL, NUSAP, PBA. 2a: EE, FDC, NUSAP, SI. 2b: EE, NUSAP, PBA. 3a: EE, PBA, UF. 3b: EE, PBA, UF.
101	Risk	Evaluate risk levels	26. Assess confidence in the risk levels using…	Experimental evidence	7.0 (Ig) *P*=0.02	Co	1: Ext 2a: Dat 2b: Var	1: EE, PBA, UF. 2a: EE, FDC, FL, NUSAP, PBA. 2b: EE, PBA, UF.
76	Effects	Analyse the stressor-response relationship	19. Determine the test dose for the…	Frequency	6.0 (Sc) *P*=0.89	Co	1: Var 2: Mod	1: EE, PBA, UF. 2: BBN, CI, EE, EP, PBA, ScA.
87	Effects	Integrate multiple LOEs using…	21. (Integrate multiple LOEs using…)	Quantitative methods	6.0 (Sc) *P*=0.68	Co	1: Mod 2: Dat 3: Var	1: BBN, CI, EE, EP, PBA, ScA. 2: EE, FDC, FL, PBA, SeA, ScA. 3: EE, PBA, UF.
96	Risk	Estimate and aggregate risk	25. Aggregate risk estimates for…	Assessment endpoints	6.0 (Sc) *P*=0.65	Co	1a: Ext 1b: Mod 2: Var 3: Sys	1a: EE, PBA, UF. 1b: BBN, CI, EE, EP, PBA, ScA. 2: EE, PBA, UF. 3: BBN, EE, FDC, ScA, SI.
97	Risk	Estimate and aggregate risk	25. Aggregate risk estimates for…	Stressors	6.0 (Sc) *P*=0.55	Co	1: Mod 2a: Sys 2b: Ext	1: BBN, CI, EE, EP, PBA, ScA. 2a: BBN, EE, FDC, ScA, SI. 2b: EE, PBA, UF.

							3: Var	3: EE, PBA, UF.
90	Effects	Create stressor-response profile using…	22. Single point methods showing…	Effects levels	6.0 (Sc) *P*=0.46	Co	1a: Ext 1b: Mod 2a: Dat 2b: Var	1a: EE, PBA, UF. 1b: BBN, CI, EE, EP, PBA, ScA. 2a: EE, FDC, FL, PBA, SA, ScA. 2b: EE, PBA, UF.
94	Risk	Estimate and aggregate risk	24. Estimate risk using…	Single-point profiles	6.0 (Sc) *P*=0.32	Co	1a: Ext 1b: Mod 2: Var 3: Dat	1a: EE, PBA, UF. 1b: BBN, CI, EE, EP, PBA, ScA. 2: EE, PBA, UF. 3: EE, FDC, FL, PBA, SeA, ScA.
24	Problem	Define the conceptual model	5. Consider the appropriateness of the endpoints	Relative importance of endpoints to each other	6.0 (Sc) *P*=0.26	Co	1a: Sys 1b: Mod 2a: Var 2b: Ext 3: Dat	1a: BBN, EE, FDC, ScA, SI. 1b: BBN, CI, EE, EP, PBA, ScA. 2a: EE, PBA, UF. 2b: EE, PBA, UF. 3: EE, FDC, FL, PBA, SeA, ScA.
89	Effects	Create stressor-response profile using…	22. Single point methods showing…	Extreme toxicity	6.0 (Sc) *P*=0.23	Co	1a: Dat 1b: Mod 2: Var 3: Ext	1a: EE, FDC, FL, PBA, SeA, ScA. 1b: BBN, CI, EE, EP, PBA, ScA. 2: EE, PBA, UF. 3: EE, PBA, UF.

[a] Ig=Recognised ignorance; Sc=Scenario uncertainty. Statistical significance (*P*) is used to rank like values.

[b] Co=Combined.

[c] Dat=Data; Sys=System; Var=Variability; Ext=Extrapolation; Mod=Model. Median occurrence rates are used to rank like values.

[d] BBN - Bayesian Belief Network; CI - Confidence intervals; EP - Error propagation; EE - Expert elicitation; FDC - Further data collection; FL - Fuzzy logic; LHS - Latin hypercube sampling; MCS - Monte-Carlo simulation; NUSAP - Numeral, unit, spread, assessment, and pedigree; PBA - Probability bounds analysis; PDF - Probability density function; ScA - Scenario analysis; SeA - Sensitivity analysis; SI - Stakeholder involvement; UF - Uncertainty factor.

The same approach can be followed by analysts in assigning UMTs to the remaining 77 ERA tasks in UnISERA (listed in Appendix Q), or, alternatively, to assign UMTs to the distinct groups of ERA tasks or ERA phases. Several such associations have been performed in the electronic version of UnISERA (Supplementary Material F).

7.4 Limitations

7.4.1 The uncertainty typology

The novel uncertainty typology introduced in Chapter 4 was based on the analysis of uncertainties within 171 environmental WOE assessments. There are potential limitations associated with the method used to construct the typology and its resulting categorisations.

Dependence on existing assessments to contain reliable information

Since the typology was constructed using information within existing published materials, its reliability was directly related to the reliability of those materials. This limitation may have been realised where incorrect information was presented within the sourced materials, though the peer-review process was expected to resolve these errors. Perhaps of more concern was the potential omission (rather than incorrect inclusion) of important uncertainties; key uncertainties that went unidentified in the source materials could not feature in the typology. However, the evidence base of 171 assessments was considered extensive enough to account for all potential uncertainties.

Subjectivity in the information clustering process

The clustering process used to form categorisations within and between the different types of uncertainty, whilst efficient and effective, did require an element of subjectivity on the part of the researcher. This type of qualitative clustering has the potential to blur definitions, thereby reducing the clarity of the clustered output. This potential limitation was managed as far as possible by making the clustering process transparent (see Figure 4.1).

Representativeness of the typology for application to ERAs

WOE assessments were chosen as the focus for the typology since they are a manageable subset of ERAs. However, limiting the potential breadth of studies to one ERA type may have led to biases within the evidence base, which would have been transferred into the typology. One potential bias was a focus on risk domains in which WOE assessments are commonly used. This potential limitation of the evidence base could have resulted in a lack of representativeness when applying the typology to non-WOE ERA scenarios. However, when weighed against other viable alternatives, such as using ERAs based in specific risk domains to build the evidence base, the WOE approach was deemed to be the most representative for future application of the typology.

7.4.2 The output from UnISERA

Potentially influential aspects of the method used to construct UnISERA were discussed in Section 5.8.4. Some potential limitations also exist regarding the way in which the output from UnISERA may be applied by analysts.

The scope and structure of the output

UnISERA was designed to be used for ERAs involving man-made sources of potential harm only, and cannot, at present, be applied to naturally-occurring risks such as climate change or flooding. The method on which UnISERA was built could easily be reapplied to such areas. Furthermore, the output is organised according to the detailed aspects of current ERA guidelines. Should these guidelines change significantly at any point, the output and related observations may lose some value. In this event, UnISERA would require updating to reflect the developments in relevant ERA guidance.

Using the output to inform selection of appropriate uncertainty management techniques

In order to keep the expert elicitations in the three case studies and one validation study to a realistic length, it was necessary to omit all sub-categories of uncertainty within the location dimension. Therefore, UnISERA does not account for any of the 20 sub-locations of uncertainty, just the seven locations themselves. This places a requirement on the analyst to understand the differences between the sub-locations within each location category – the difference between interspecies, intraspecies, laboratory, quantity, spatial and temporal extrapolation, for example – and to make this distinction in their ERAs. This understanding is important because it may influence the appropriateness of any UMTs selected. The enforced lack of granularity in the output of UnISERA is deemed the most significant limitation of this research.

7.5 Future research

7.5.1 The uncertainty typology

The stated potential limitations within the uncertainty typology introduced in this research are reflective of a lack of formal validation performed on its contents, rather than of its reliability. Therefore, the potential for work in this area is driven by the need to compare the defined uncertainty categorisations to those in other similarly compiled typologies. Since no other typologies exist (in the context of environmental risk) which follow the same extensive evidence-based method, a direct comparison is neither appropriate nor possible. Validation of the uncertainty typology could therefore be obtained in two ways:

- the clustering process could be repeated, using the same WOE evidence base but with either a different researcher performing the clustering, or using software that implements cluster analysis algorithms;
- an alternative evidence base could be compiled, using the same formalised literature search method, to contain the full available set of ERAs in which uncertainty analysis (or a component of uncertainty analysis) features. This (probably enormous) set of articles could then be sampled randomly to produce a manageable subset (e.g. 171) for analysis.

The resulting set of uncertainty characterisations could then be compared to those in the current typology.

7.5.2 UnISERA

Identifying the aspects of the output from UnISERA that disagreed with the Validation Case Study (see Chapter 6) offers opportunity for immediate modification, which could occur in three forms:

- expansion of the existing system, through the addition of extra case studies;
- additional validation, by following the method described in Chapter 5 and either using the same three case studies with alternative experts, or using new case studies and new experts to attempt to recreate the output from UnISERA;
- application of the existing or modified version of UnISERA to ERA scenarios by risk analysts.

Any or all of these three options could realistically be applied, according to temporal and financial constraints.

Any future development of the UnISERA approach should also attempt to incorporate the sub-locations of uncertainty, and to associate them with the defined ERA tasks. This would increase the granularity of the guidance, further aiding the analyst.

One of the biggest restrictions concerning the future application of the output from UnISERA is its focus on man-made sources of risk only. A key task could therefore be to create a companion version for use with naturally-occurring sources of potential harm, which conforms to relevant ERA structures.

7.5.3 Beyond the typology and UnISERA

In advancing understanding associated with the characterisation and identification of uncertainties in ERAs this research has also revealed several other knowledge gaps, which may be addressed through the following research directions and questions:

- *The appropriateness and performance of uncertainty management techniques.* Can the existing set of UMTs be used to manage the combinations of levels, natures, and locations (and sub-locations) of uncertainty described in this research, and which UMTs perform best? A sensitivity analysis approach could be adopted to test the capabilities of UMTs in different ERA scenarios.

- *The influence of uncertainty on environmental risk management.* Suppose two ERAs exist in a chosen risk domain: one with a full uncertainty analysis and one without any formal treatment of uncertainty. Does the uncertainty analysis process improve the basis for implementing risk management actions? Application of these two types of ERAs by risk managers, and observations regarding the actions that they advise, would allow for the influence of uncertainty to be assessed qualitatively.
- *Reliability of controlling regulation and guidance.* Do the relevant regulatory tools enforce a thorough treatment of uncertainty in ERAs in different risk domains, and are guidance documents providing analysts with the right information? This could be achieved through the analysis of ERAs submitted to regulatory bodies (as part of a regulated process), and comparison to the research in this thesis.
- *Regulating for uncertainty in emerging risk domains.* How does regulation differ in its approach to uncertainty in established risk domains (e.g. pesticides) compared with novel or emerging risk domains (e.g. engineered nanomaterials)? Is regulation keeping up with scientific advancement? A case study approach involving regulatory analysis (across different bodies) and expert engagement could be implemented.

7.6 A summary of the practical applications of this research

The research presented in this thesis can be applied in an attempt to move uncertainty characterisation and identification within ERAs forward. Specifically, ERA analysts who have access to this research will be equipped with:

- a single typology of uncertainties with which to understand the potential types of uncertainties that can exist within ERAs (see Chapter 4); and
- validated guidance for identifying uncertainties within the tasks, groups of tasks, and phases of ERAs (see Chapters 5 and 6).

This latter point may also reduce the previously existing sole reliance on the analyst to identify uncertainties within ERAs, before being able to implement one or more appropriate techniques to manage those uncertainties.

This research will be distributed through publication in peer-reviewed journals (see page xvi), with the electronic version of UnISERA made available for download, and potentially within future relevant environmental risk assessment guidance documents.

References

Abrantes N., R. Pereira, and F. Gonçalves. 2010. Occurrence of Pesticides in Water, Sediments, and Fish Tissues in a Lake Surrounded by Agricultural Lands: Concerning Risks to Humans and Ecological Receptors. *Water, Air and Soil Pollution* 212: 77-88.

Acosta H., D. Wu, and B. Forrest. 2010. Fuzzy experts on recreational vessels, a risk modelling approach for marine invasions. *Ecological Modelling* 221(5): 850-863.

Agüero A., P. Pinedo, I. Simón, D. Cancio, M. Moraleda, C. Trueba, and D. Pérez-Sánchez. 2008. Application of the Spanish methodological approach for biosphere assessment to a generic high-level waste disposal site. *Science of the Total Environment* 403(1-3): 34-58.

Ahlers J., F. Stock, and B. Werschkun. 2008. Integrated testing and intelligent assessment - New challenges under REACH. *Environmental Science and Pollution Research* 15(7): 565-572.

Alcamo J. and J. Bartnicki. 1987. A framework for error analysis of a long-range transport model with emphasis on parameter uncertainty. *Atmospheric Environment (1967)* 21(10): 2121-2131.

Alden III R.W., L.W. Hall Jr., D.M. Dauer, and D.T. Burton. 2005. An integrated case study for evaluating the impacts of an oil refinery effluent on aquatic biota in the Delaware River: Integration and analysis of study components. *Human and Ecological Risk Assessment* 11(4): 879-936.

Allen R.W., M.H. Criqui, A.V. Diez Roux, M. Allison, S. Shea, R. Detrano, L. Sheppard, N.D. Wong, K.H. Stukovsky, and J.D. Kaufman. 2009. Fine particulate matter air pollution, proximity to traffic, and aortic atherosclerosis. *Epidemiology* 20(2): 254-264.

Almeida S.M., M.M. Farinha, M.G. Ventura, C.A. Pio, M.C. Freitas, and M.A Reis. 2007. Measuring Air Particulate Matter in large urban areas for health effect assessment. *Water, Air and Soil Pollution* 179: 43-55.

Altieri M.A. 2000. The ecological impacts of transgenic crops on agroecosystem health. *Ecosystem Health* 6(1): 13-23.

Alvarez-Guerra M., J. Viguri, and N. Voulvoulis. 2009. A multicriteria-based methodology for site prioritisation in sediment management. *Environment International* 35(6): 920-930.

An Y.J., C.G. Yoon, K.W. Jung, and J.H. Ham. 2007. Estimating the microbial risk of E. coli in reclaimed wastewater irrigation on paddy field. *Environmental Monitoring and Assessment* 129(1-3): 53-60.

Anderson P.L., R.L. Hellmich, J.R. Prasifka, L.C. Lewis. 2005. Effects on fitness and behavior of monarch butterfly larvae exposed to a combination of Cry1Ab-expressing corn anthers and pollen. *Environmental Entomology* 34(4): 944-952.

Apitz S.E., A. Barnati, A.G. Bernstein, M. Bocci, E. Delaney, and L. Montobbio. 2007. The assessment of sediment screening risk in Venice Lagoon and other coastal areas using international sediment quality guidelines. *Journal of Soils and Sediments* 7(5): 326-341.

ApSimon H.M., R.F. Warren, and S. Kayin. 2002. Addressing uncertainty in environmental modelling: A case study of integrated assessment of strategies to combat long-range transboundary air pollution. *Atmospheric Environment* 36(35): 5417-5426.

Arhonditsis G., S. Qian, C. Stow, E. Lamon, and K. Reckhow. 2007. Eutrophication risk assessment using Bayesian calibration of process-based models: Application to a mesotrophic lake. *Ecological Modelling* 208(2-4): 215-229.

Aschberger K., C. Micheletti, B. Sokull-Kluttgen, and F.M. Christensen. 2011. Analysis of currently available data for characterising the risk of engineered nanomaterials to the environment and human health – lessons learned from four case studies. *Environment International* 37(6): 1143-1156.

Ascough II J., H. Maier, J. Ravalico, and M. Strudley. 2008. Future research challenges for incorporation of uncertainty in environmental and ecological decision-making. *Ecological Modelling* 219(3-4): 383-399.

Aspinall W., G. Woo, B. Voight, and P.J. Baxter. 2003. Evidence-based volcanology: application to eruption crises. *Journal of Volcanology and Geothermal Research* 128: 273-285.

van Asselt M.B.A. and J.P. Rotmans. 2002. Uncertainty in integrated assessment modelling. *Climatic Change* 54(1): 75-105.

Auer C. 2008. Ecological risk assessment and regulation for genetically-modified ornamental plants. *Critical Reviews in Plant Sciences* 27: 255-271.

Avagliano S., L. Parrella. 2009. Managing uncertainty in risk-based corrective action design: Global sensitivity analysis of contaminant fate and exposure models used in the dose assessment. *Environmental Modeling and Assessment* 141: 47-57.

Avagliano S., A. Vecchio, and V. Belgiorno. 2005. Sensitive parameters in predicting exposure contaminants concentration in a risk assessment process. *Environmental Monitoring and Assessment* 111(1-3): 133-148.

Babendreier J.E. and K.J. Castleton. 2005. Investigating uncertainty and sensitivity in integrated, multimedia environmental models: Tools for FRAMES-3MRA. *Environmental Modelling and Software* 20(8): 1043-1055.

Baccou J., E. Chojnacki, C. Mercat-Rommens, and C. Baudrit. 2008. Extending Monte Carlo simulations to represent and propagate uncertainties in presence of incomplete knowledge: Application to the transfer of a radionuclide in the environment. *Journal of Environmental Engineering* 134(5): 362-368.

Barnthouse L. 2008. The strengths of the ecological risk assessment process: Linking science to decision making. *Integrated environmental assessment and management* 4: 299-305.

Barron M.G., S.E. Duvall, and K.J. Barron. 2004. Retrospective and current risks of mercury to panthers in the Florida everglades. *Ecotoxicology* 13(3): 223-229.

Batley G.E., G.A. Burton, P.M. Chapman, and V.E. Forbes. 2002. Uncertainties in sediment quality weight-of-evidence (WOE) assessments. *Human and Ecological Risk Assessment* 8(7): 1517-1547.

Batzias F. and C.G. Siontorou. 2007. A novel system for environmental monitoring through a cooperative/synergistic scheme between bioindicators and biosensors. *Journal of Environmental Management* 82(2): 221-239.

Baudrit C., D. Guyonnet, and D. Dubois. 2007. Joint propagation of variability and imprecision in assessing the risk of groundwater contamination. *Journal of Contaminant Hydrology* 93(1-4): 72-84.

Beck M. 1987. Water quality modeling: a review of the analysis of uncertainty. *Water Resources Research* 23(8): 1393-1442.

Bedford T. and R. Cooke. 2001. *Probabilistic risk analysis: foundations and methods*. Cambridge: Cambridge University Press.

Benekos I.D., C.A. Shoemaker, and J.R. Stedinger. 2007. Probabilistic risk and uncertainty analysis for bioremediation of four chlorinated ethenes in groundwater. *Stochastic Environmental Research and Risk Assessment* 21(4): 375-390.

Benke K.K. and A.J. Hamilton. 2008. Quantitative microbial risk assessment: Uncertainty and measures of central tendency for skewed distributions. *Stochastic Environmental Research and Risk Assessment* 22(4): 533-539.

Bennett J.R., C.A. Kaufman, I. Koch, J. Sova, and K.J. Reimer. 2007. Ecological risk assessment of lead contamination at rifle and pistol ranges using techniques to account for site characteristics. *Science of the Total Environment* 374(1): 91-101.

Beer T. 2006. Ecological risk assessment and quantitative consequence analysis. *Human and Ecological Risk Assessment* 12(1): 51-65.

van den Belt H. 2003. Debating the Precautionary Principle: "guilty until proven innocent" or "innocent until proven guilty"? *Plant Physiology* 132(3): 1122-1126.

Betha R. and R. Balasubramanian. 2011. Emissions of particulate-bound elements from stationary diesel engine: Characterization and risk assessment. *Atmospheric Environment* 45(30): 5273-5281.

Bevington P. and D. Robinson. 2002.*Data reduction and error analysis for the physical sciences*, Vol. 336. New York: McGraw-Hill.

Beyer C., S. Altfelder, W.H.M. Duijnisveld, and T. Streck. 2009. Modelling spatial variability and uncertainty of cadmium leaching to groundwater in an urban region. *Journal of Hydrology* 369(3-4): 274-283.

Biskos G. and A. Schmidt-Ott. 2012. Airborne Engineered Nanoparticles: Potential Risks and Monitoring Challenges for Assessing their Impacts on Children. *Paediatric Respiratory Reviews* 13(2): 79-83.

Bittueva M.M., S.K. Abilev, and V.A. Tarasov. 2007. Efficiency of the prediction of carcinogenic activities of chemical substances based on scoring somatic mutations in the soybean Glycine max (L.) Merrill. *Russian Journal of Genetics* 43(1): 64-72.

Blazkova S., and K. Beven. 2004. Flood frequency estimation by continuous simulation of subcatchment rainfalls and discharges with the aim of improving dam safety assessment in a large basin in the Czech Republic. *Journal of Hydrology* 292(1-4): 153-172.

Boldo E., C. Linares, J. Lumbreras, R. Borge, A. Narros, J. García-Pérez, P. Fernández-Navarro, B. Pérez-Gómez, N. Aragonés, R. Ramis, *et al*. 2011. Health impact assessment of a reduction in ambient PM2.5 levels in Spain. *Environment International* 37(2): 342-348.

Borsuk M.E., P. Reichert, A. Peter, E. Schager, and P. Burkhardt-Holm. 2006. Assessing the decline of brown trout (Salmo trutta) in Swiss rivers using a Bayesian probability network. *Ecological Modelling* 192(1-2): 224-244.

Bosgra S., H.v.d. Voet, P.E. Boon, and W. Slob. 2009. An integrated probabilistic framework for cumulative risk assessment of common mechanism chemicals in food: An example with organophosphorus pesticides. *Regulatory Toxicology and Pharmacology* 54(2): 124-133.

Bosgra S., P.M.J. Bos, T.G. Vermeire, R.J. Luit, and W. Slob. 2005. Probabilistic risk characterization: An example with di(2-ethylhexyl) phthalate. *Regulatory Toxicology and Pharmacology* 43(1): 104-113.

Bowden R. 2004. Building confidence in geological models. *Geological Society, London, Special Publications* 239(1): 157.

Boxall A.B.A., Q. Chaudhry, C. Sinclair, A. Jones, R. Aitken, B. Jefferson, and C. Watt. 2007. *Current and future predicted environmental exposure to engineered nanoparticles*. Central Science Laboratory Report, University of York, prepared for the UK Department of Environment, Food and Rural Affairs.

Brèchignac F. and M. Doi. 2009. Challenging the current strategy of radiological protection of the environment: arguments for an ecosystem approach. *Journal of Environmental Radioactivity* 100(12): 1125-1134.

Briggs D.J. 2008. A framework for integrated environmental health impact assessment of systemic risks. *Environmental Health* 7(61).

van den Brink P.J., S.J.H. Crum, R. Gylstra, F. Bransen, J.G.M. Cuppen, T.C.M. Brock. 2009. Effects of a herbicide-insecticide mixture in freshwater microcosms: Risk assessment and ecological effect chain. *Environmental Pollution* 157(1): 237-249.

van den Brink C., W.J. Zaadnoordijk, S. Burgers, and J. Griffioen. 2008. Stochastic uncertainties and sensitivities of a regional-scale transport model of nitrate in groundwater. *Journal of Hydrology* 361(3-4): 309-318.

British Standards Institution. 2007. *PAS 136: Terminology for Nanomaterials*. London: British Standards Institution.

Brook R.D., H.H. Shin, R.L. Bard, R.T. Burnett, A. Vette, C. Croghan, J. Thornburg, C. Rodes, and R. Williams. 2011. Exploration of the rapid effects of personal fine particulate matter exposure on arterial hemodynamics and vascular function during the same day. *Environmental Health Perspectives* 119(5): 688-694.

Brouwer R. and C. De Blois. 2008. Integrated modelling of risk and uncertainty underlying the cost and effectiveness of water quality measures. *Environmental Modelling & Software* 23(7): 922-937.

Brown B.B., S.W. Cochran, and N.C. Dalkey. 1969. *The Delphi method, II: structure of experiments*. Rand Report RM-5957-PR. Santa-Monica: Rand Corporation.

Brown J. 2004. Knowledge, uncertainty and physical geography: towards the development of methodologies for questioning belief. *Transactions of the Institute of British Geographers* 29(3): 367-381.

Buekers J., E. Steen Redeker, and E. Smolders. 2009. Lead toxicity to wildlife: Derivation of a critical blood concentration for wildlife monitoring based on literature data. *Science of the Total Environment* 407(11): 3431-3438.

Burgert S., R.B. Schäfer, K. Foit, M. Kattwinkel, L. Metzeling, R. MacEwan, B.J. Kefford, M. Liess. 2011. Modelling aquatic exposure and effects of insecticides - Application to south-eastern Australia. *Science of the Total Environment* 409(14): 2807-2814.

Burgman M.A., D.A. Keith, and T.V. Walshe. 1999. Uncertainty in comparative risk analysis for threatened Australian plant species. *Risk Analysis* 19(4): 585-598.

Burton Jr. G.A., M.S. Greenberg, C.D. Rowland, C.A. Irvine, D.R. Lavoie, J.A. Brooker, L. Moore, D.F.N. Raymer, and R.A. McWilliam. 2005. In situ exposures using caged organisms: A multi-compartment approach to detect aquatic toxicity and bioaccumulation. *Environmental Pollution* 134(1): 133-144.

Burton Jr. G.A., P.M. Chapman, and E.P. Smith. 2002. Weight-of-evidence approaches for assessing ecosystem impairment. *Human and Ecological Risk Assessment* 8(7): 1657-1673.

Calabrese E.J., L.A. Baldwin, P.T. Kostecki, and T.L. Potter. 1997. A toxicologically based weight-of-evidence methodology for the relative ranking of chemicals of endocrine disruption potential. *Regulatory Toxicology and Pharmacology* 261I: 36-40.

Caley P., W.M. Lonsdale, and P.C. Pheloung. 2006. Quantifying uncertainty in predictions of invasiveness. *Biological Invasions* 8(2): 277-286.

Campbell J.E. and D.E. Longsine. 1990. Application of generic risk assessment software to radioactive waste disposal. *Reliability Engineering and System Safety* 30(1-3): 183-193.

Cañellas-Boltà S., R. Strand, and B. Killie. 2005. Management of environmental uncertainty in maintenance dredging of polluted harbours in Norway. *Water Science and Technology* 52(6): 93-98.

Carlon C., L. Pizzol, A. Critto, and A. Marcomini. 2008. A spatial risk assessment methodology to support the remediation of contaminated land. *Environment International* 34(3): 397-411.

Carrington C.D., G.M. Cramer, and P.M. Bolger. 1997. A risk assessment for methylmercury in tuna. *Water, Air, and Soil Pollution* 97(3-4): 273-283.

Cesar A., A. Marín, L. Marin-Guirao, R. Vita, J. Lloret, and T.A. Del Valls.2009. Integrative ecotoxicological assessment of sediment in Portmán Bay (southeast Spain). *Ecotoxicology and Environmental Safety* 72: 1832-1841.

Chapman PM. 2007. Determining when contamination is pollution - Weight of evidence determinations for sediments and effluents. *Environment International* 334: 492-501.

Chen. T.H., C.Y. Lin, and M.C. Tseng. 2011. Behavioral effects of titanium dioxide nanoparticles on larval zebrafish (Danio rerio). *Marine Pollution Bulletin* 63(5-12): 303-308.

Chen Z., A.M. Yadghar, L. Zhao, and Z. Mi. 2011. A review of environmental effects and management of nanomaterials. *Toxicological & Environmental Chemistry* 93: 1227-50.

Chen C., H. Ma, and K. Reckhow. 2007. Assessment of water quality management with a systematic qualitative uncertainty analysis. *Science of the Total Environment* 374(1): 13-25.

Chen Y.C. and H.W. Ma. 2007. Combining the cost of reducing uncertainty with the selection of risk assessment models for remediation decision of site contamination. *Journal of Hazardous Materials* 141(1): 17-26.

Chèvre N., P. Edder, D. Ortelli, E. Tatti, S. Erkman, and F. Rapin. 2008. Risk assessment of herbicide mixtures in a large European lake. *Environmental Toxicology* 23: 269-277.

Chèvre N., C. Loepfe, H. Singer, C. Stamm, K. Fenner, and B.I. Escher. 2006. Including mixtures in the determination of water quality criteria for herbicides in surface water. *Environmental Science and Technology* 15(40): 426-35.

Chio C.P., W.Y. Chen, W.C. Chou, N.H. Hsieh, M.P. Ling, and C.M. Liao. 2012. Assessing the potential risks to zebrafish posed by environmentally relevant copper and silver nanoparticles. *Science of the Total Environment* 420: 111-118.

Chowdhury S., P. Champagne, and P.J. McLellan. 2009. Uncertainty characterization approaches for risk assessment of DBPs in drinking water: A review. *Journal of Environmental Management* 90(5): 1680-1691.

Chowdhury R. and P. Flentje. 2003. Role of slope reliability analysis in landslide risk management. *Bulletin of Engineering Geology and the Environment* 62(1): 41-46.

Ciffroy P., G. Durrie, and J.M. Garnier. 2009. Probabilistic distribution coefficients (K(d)s) in freshwater for radioisotopes of Ag, Am, Ba, Be, Ce, Co, Cs, I, Mn, Pu, Ra, Ru, Sb, Sr and Th - Implications for uncertainty analysis of models simulating the transport of radionuclides in rivers. *Journal of Environmental Radioactivity* 100(9): 785-794.

Clemen R. and R. Winkler. 1999. Combining probability distributions from experts in risk analysis. *Risk Analysis* 19: 187-203.

Cohen I.B. and A. Whitman. 1999. *Isaac Newton: The Principia, Mathematical principles of natural philosophy, a new translation*. Richmond (CA): University of California Press.

Collins J.F., G.V. Alexeeff, D.C. Lewis, D.E. Dodge, M.A. Marty, T.R. Parker, J.D. Budroe, R.H.F. Lam, M.J. Lipsett, J.R. Fowles, *et al.* 2004. Development of acute inhalation reference exposure levels (RELs) to protect the public from predictable excursions of airborne toxicants. *Journal of Applied Toxicology* 24(2): 155-166.

Collins G., J.N. Kremer, and I. Valiela. 2000. Assessing uncertainty in estimates of nitrogen loading to estuaries for research, planning, and risk assessment. *Environmental Management* 25(6): 635-645.

Cooke R.M. 1991. *Experts in Uncertainty; Opinion and Subjective Probability in Science*. New York & Oxford: Oxford University Press.

Cooke R.M. and K.N. Probst. 2006. *Highlights of the Expert Judgment Policy Symposium and Technical Workshop*. Conference Summary.

Costanza R., S. Funtowicz, and J. Ravetz. 1992. Assessing and communicating data quality in policy-relevant research. *Environmental Management* 16(1): 121-131.

Cothern C.R., W.A. Coniglio, and W.L. Marcus. 1986. Development of quantitative estimates of uncertainty in environmental risk assessments when the scientific data base is inadequate. *Environment International* 12(6): 643-647.

Crane J.L. and D.D. MacDonald. 2003. Applications of Numerical Sediment Quality Targets for Assessing Sediment Quality Conditions in a US Great Lakes Area of Concern. *Environmental Management* 32(1): 128-140.

Critto A., S. Torresan, E. Semenzin, S. Giove, M. Mesman, A.J. Schouten, M. Rutgers, and A. Marcomini. 2007. Development of a site-specific ecological risk assessment for contaminated sites: Part I. A multi-criteria based system for the selection of ecotoxicological tests and ecological observations. *Science of the Total Environment* 3791: 16-33.

Croke B., J. Ticehurst, R. Letcher, J. Norton, L. Newham, and A. Jakeman. 2007. Integrated assessment of water resources: Australian experiences. *Water resources management* 21(1): 351-373.

Culp J.M., C.L. Podemski, K.J. Cash, and R.B. Lowell. 2000. A research strategy for using stream microcosms in ecotoxicology: Integrating experiments at different levels of biological organization with field data. *Journal of Aquatic Ecosystem Stress and Recovery* 7(2): 167-176.

Cupit M., O. Larsson, C. De Meeûs, G.H. Eduljee, and M. Hutton. 2002. Assessment and management of risks arising from exposure to cadmium in fertilisers - II. *Science of the Total Environment* 291(1-3): 189-206.

Cuppen J.G.M., P.J. Van Den Brink, E. Camps, K.F. Uil, and T.C.M. Brock. 2000. Impact of the fungicide carbendazim in freshwater microcosms. I. Water quality, breakdown of particulate organic matter and responses of macroinvertebrates. *Aquatic Toxicology* 48(2-3): 233-250.

Dale V., G. Biddinger, M. Newman, J. Oris, G. Suter, T. Thompson, T. Armitage, J. Meyer, R. Allen-King, G. Burton, *et al*. 2008. Enhancing the ecological risk assessment process. *Integrated environmental assessment and management* 4(3): 306-313.

Dale V., J. Meyer, and United States Environmental Protection Agency Science Advisory Board. 2007. *Advice to EPA on advancing the science and application of ecological risk assessment in environmental decision making: a report of the U.S. EPA Science*

Advisory Board. Washington: United States Environmental Protection Agency, Office of the Administrator.

Dalkey N.C. 1969. *The Delphi method: an experimental study of group opinion*. Rand Report RM-5888-PR. Santa-Monica: Rand Corporation.

Damalas C.A. and I.G. Eleftherohorinos. 2011. Pesticide exposure, safety issues, and risk assessment indicators. *International Journal of Environmental Research and Public Health* 8: 1402-1419.

Damásio J., D. Barceló, R. Brix, C. Postigo, M. Gros, M. Petrovic, S. Sabater, H. Guasch, M.L. de Alda, and C. Barata. 2011. Are pharmaceuticals more harmful than other pollutants to aquatic invertebrate species: A hypothesis tested using multi-biomarker and multi-species responses in field collected and transplanted organisms. *Chemosphere* 85(10): 1548-1554.

Daniels J.I., K.T. Bogen, and L.C. Hall. 2000. Analysis of uncertainty and variability in exposure to characterize risk: Case study involving trichloroethylene groundwater contamination at Beale Air Force Base in California. *Water, Air, and Soil Pollution* 123(1-4): 273-298.

Darbra R.M., E. Eljarrat, and D. Barcelo. 2008. How to measure uncertainties in environmental risk assessment. *Trends in Analytical Chemistry* 27(4): 377-385.

Deck L.B., E.S. Post, E. Smith, M. Wiener, K. Cunningham, and H. Richmond. 2001. Estimates of the health risk reductions associated with attainment of alternative particulate matter standards in two U.S. cities. *Risk Analysis* 21(5): 821-836.

DelValls T.A., and I. Riba. 2007. A weight of evidence approach to assess sediment quality in the Guadalquivir estuary. *Aquatic Ecosystem Health and Management* 10(1): 101-106.

Department for Environment, Food and Rural Affairs. 2011. *Guidelines for environmental risk assessment and management: Green Leaves III*. London: HMSO.

Department of Environment, Transport and the Regions/Environment Agency and Institute for Environmental Health. 2000. *Guidelines for environmental risk assessment and management: revised departmental guidance*. London: HMSO.

Department of Health and Ageing. 2002. *Environmental Health Risk assessment: guidelines for assessing human health risks from environmental hazards*. Canberra: HMSO.

Dewulf A., M. Craps, R. Bouwen, T. Taillieu, C. Pahl-Wostl, and M. Krayer von Krauss. 2005. Integrated management of natural resources: dealing with ambiguous issues, multiple actors and diverging frames. *Water Science and Technology* 52(6): 115-124.

Dey W.P., S.M. Jinks, and G.J. Lauer. 2000. The 316(b) assessment process: Evolution towards a risk-based approach. *Environmental Science and Policy 3Suppl*.

Diaconis P., S. Holmes, and R. Montgomery. 2007. Dynamical Bias in the Coin Toss. *Society for Industrial and Applied Mathematics (SIAM) Review* 49(2): 211-235.

Diaconis P. 1998. A Place for Philosophy? The Rise of Modeling in Statistics. *Quarterly Journal of Applied Mathematics* 56(4): 797-805.

Diodato N. and M. Ceccarelli. 2005. Environinformatics in ecological risk assessment of agroecosystems pollutant leaching. *Stochastic Environmental Research and Risk Assessment* 19(4): 292-300.

Díaz R.V. and E. Rosa Dominguez. 2009. Health risk by inhalation of PM2.5 in the metropolitan zone of the City of Mexico. *Ecotoxicology and Environmental Safety* 72(3): 866-871.

Dively G.P., R. Rose, M.K. Sears, R.L. Hellmich, D.E. Stanley-Horn, D.D. Calvin, J.M. Russo, and P.L. Anderson. 2004. Effects on monarch butterfly larvae (Lepidoptera: Danaidae) after continuous exposure to Cry1Ab-expressing corn during anthesis. *Environmental Entomology* 33(4): 1116-1125.

Doull J., K.K. Rozman, and M.C. Lowe. 1996. Hazard evaluation in risk assessment: whatever happened to sound scientific judgment and weight of evidence? *Drug Metabolism Reviews* 28: 285-299.

Ducey M. and B. Larson. 1999. A fuzzy set approach to the problem of sustainability. *Forest Ecology and Management* 115(1): 29-40.

Dunham J.B., M.K. Young, R.E. Gresswell, and B.E. Rieman. 2003. Effects of fire on fish populations: Landscape perspectives on persistence of native fishes and nonnative fish invasions. *Forest Ecology and Management* 178(1-2): 183-196.

Dussault É.B., V.K. Balakrishnan, E. Sverko, K.R. Solomon, and P.K. Sibley. 2008. Toxicity of human pharmaceuticals and personal care products to benthic invertebrates. *Environmental Toxicology and Chemistry* 27(2): 425-432.

Eckelman M.J., M.S. Mauter, J.A. Isaacs, and M. Elimelech. 2012. New perspectives on nanomaterial aquatic ecotoxicity: Production impacts exceed direct exposure impacts for carbon nanotoubes. *Environmental Science and Technology* 46(5): 2902-2910.

Echevarria G., M.I. Sheppard, and J. Morel. 2001. Effect of pH on the sorption of uranium in soils. *Journal of Environmental Radioactivity* 53(2): 257-264.

Eckhardt R. 1987. Stan Ulam, John von Neumann, and the Monte Carlo method. *Los Alamos Science, Special Issue* 15: 131-137.

Efroymson R.A., D.S. Jones, and A.J. Gold. 2007. An ecological risk assessment framework for effects of onsite wastewater treatment systems and other localized sources of nutrients on aquatic ecosystems. *Human and Ecological Risk Assessment* 13(3): 574-614.

El-Ghonemy H., L. Watts, and L. Fowler. 2005. Treatment of uncertainty and developing conceptual models for environmental risk assessments and radioactive waste disposal safety cases. *Environment International* 31(1): 89-97.

Enick O.V. and M.M. Moore. 2007. Assessing the assessments: Pharmaceuticals in the environment. *Environmental Impact Assessment Review* 27(8): 707-729.

Environmental Protection Act. 1990. *Environmental Protection Act: Chapter 43, Section 1*. London: The Stationery Office.

European Chemicals Agency. 2008. *Guidance on information requirements and chemical safety assessment: Chapter R.19, uncertainty analysis*. Saarbrücken (Germany): Dictus Publishing.

European Commission. 2008. *Regulatory aspects of nanomaterials*. Brussels: Communication from the Commission to the European Parliament, the Council, and the European Economic and Social Committee.

European Commission. 2001. European Commission Directive 2001/18/EC of the European Parliament and of the Council of 12 March 2001 on the deliberate release into the environment of genetically modified organisms and repealing Council Directive 90/220/EEC. *Official Journal of the European Communities* L106: 1-39.

European Environment Agency. 2007. *Late lessons from early warnings the Precautionary Principle 1896-2000: Implications for Policymaking and Research*.

European Food Safety Authority. 2010. Guidance on the environmental risk assessment of genetically modified plants. *The European Food Safety Authority (EFSA) Journal* 8(11): 1879.

European Food Safety Authority. 2006. Opinion of the scientific committee related to uncertainties in dietary exposure assessment. *The European Food Safety Authority (EFSA) Journal* 438: 1-54.

Fairman R., C. Mead, W. Williams, and A. de Medio Ambiente. 1998. *Environmental risk assessment: approaches, experiences and information sources*. Luxembourg: Office for Official Publications of the European Communities.

Falagas M.E., E.I. Pitsouni, G.A. Malietzis, and G. Pappas. 2008. Comparison of PubMed, Scopus, Web of Science, and Google Scholar: Strengths and weaknesses. *The Federation of American Societies for Experimental Biology Journal 22*(2): 338-342.

Farkas J., P. Christian, J.A. Gallego-Urrea, N. Roos, M. Hassellöv, K.E. Tollefsen, and K.V. Thomas. 2011. Uptake and effects of manufactured silver nanoparticles in rainbow trout (Oncorhynchus mykiss) gill cells. *Aquatic Toxicology* 101(1): 117-125.

Faucheux S. and G. Froger. 1995. Decision-making under environmental uncertainty. *Ecological Economics* 15(1): 29-42.

Fewtrell L., S.M. Macgill, D. Kay, and D. Casemore. 2001. Uncertainties in risk assessment for the determination of drinking water pollutant concentrations: Cryptosporidium case study. *Water Research* 35(2): 441-447.

Fiksel J. 1985. Quantitative risk analysis for toxic chemicals in the environment. *Journal of Hazardous Materials* 10(2-3): 227-240.

Filipsson M., M. Lindström, P. Peltola, and T. Öberg. 2009. Exposure to contaminated sediments during recreational activities at a public bathing place. *Journal of Hazardous Materials* 171(1-3): 200-207.

Finkel A. 1990. *Confronting Uncertainty in Risk Management: A Guide for Decision–Makers*. Washington: Center for Risk Management, Resources for the Future.

Finnveden G., M.Z. Hauschild, T. Ekvall, J. Guinée, R. Heijungs, S. Hellweg, A. Koehler, D. Pennington, and S. Suh. 2009. Recent developments in Life Cycle Assessment. *Journal of Environmental Management* 91(1): 1-21.

Fischer D.L. 2005. Accounting for differing exposure patterns between laboratory tests and the field in the assessment of long-term risks of pesticides to terrestrial vertebrates. *Ecotoxicology* 14(8): 853-862.

Fish R., M. Winter, D.M. Oliver, D. Chadwick, T. Selfa, A.L. Heathwaite, and C. Hodgson. 2009. Unruly pathogens: eliciting values for environmental risk in the context of heterogeneous expert knowledge. *Environmental Science and Policy* 12(3): 281-296.

Forbes V.E. and P. Calow. 2002. Applying weight-of-evidence in retrospective ecological risk assessment when quantitative data are limited. *Human and Ecological Risk Assessment* 8(7): 1625-1639.

Forbes V.E., P. Calow, and R.M. Sibly. 2001. Are current species extrapolation models a good basis for ecological risk assessment? *Environmental Toxicology and Chemistry* 20: 442-447.

Fuhrer J. 2009. Ozone risk for crops and pastures in present and future climates. *Naturwissenschaften* 96(2): 173-194.

Funtowicz S. and J. Ravetz. 1990.*Uncertainty and quality in science for policy*, Vol. 15. Dordrecht: Kluwer.

Gathmann A., L. Wirooks, L.A. Hothorn, D. Bartsch, and I. Schuphan. 2006. Impact of Bt maize pollen (MON810) on lepidopteran larvae living on accompanying weeds. *Molecular Ecology* 15(9): 2677-2685.

Gillund F., K. Kjшlberg, M. Krayer von Krauss, and A. Myhr. 2008. Do uncertainty analyses reveal uncertainties? Using the introduction of DNA vaccines to aquaculture as a case. *Science of the Total Environment* 407(1): 185-196.

Godduhn A. and L.K. Duffy. 2003. Multi-generation health risks of persistent organic pollution in the far north: Use of the precautionary approach in the Stockholm Convention. *Environmental Science and Policy* 64: 341-353.

Golden R.J., S.E. Holm, D.E. Robinson, P.H. Julkunen, and E.A. Reese. 1997. Chloroform mode of action: Implications for cancer risk assessment. *Regulatory Toxicology and Pharmacology* 26(2): 142-155.

González M. and A. Herrador. 2007. A practical guide to analytical method validation, including measurement uncertainty and accuracy profiles. *Trends in Analytical Chemistry* 26(3): 227-238.

Goodman J., E. McConnell, I. Sipes, R. Witorsch, T. Slayton, C. Yu, A. Lewis, and L. Rhomberg. 2006. An updated weight of the evidence evaluation of reproductive and developmental effects of low doses of bisphenol A. *Critical Reviews in Toxicology* 36(5): 387-457.

Goswami E., T. Larson, T. Lumley, and L.-J.S. Liu. 2002. Spatial characteristics of fine particulate matter: Identifying representative monitoring locations in Seattle, Washington. *Journal of the Air and Waste Management Association* 52(3): 324-333.

Gottschalk F. and B. Nowack. 2011. The release of engineered nanomaterials to the environment. *Journal of Environmental Monitoring* 13(5): 1145-1155.

Gottschalk F., T. Sonderer, R.W. Scholz, and B. Nowack. 2010. Possibilities and limitations of modeling environmental exposure to engineered nanomaterials by probabilistic material flow analysis. *Environmental Toxicology and Chemistry* 29: 1036-48.

Gottschalk F., T. Sonderer, R.W. Scholz, and B. Nowack. 2009. Modeled environmental concentrations of engineered nanomaterials (TiO2, ZnO, Ag, CNT, fullerenes) for different regions. *Environmental Science and Technology* 43(24): 9216-9222.

Greco S.L., A.M. Wilson, J.D. Spengler, and J.I. Levy. 2007. Spatial patterns of mobile source particulate matter emissions-to-exposure relationships across the United States. *Atmospheric Environment* 41(5): 1011-1025.

Greenberg M.M. 1997. The central nervous system and exposure to toluene: A risk characterization. *Environmental Research* 72(1): 1-7.

Greene N.A. and V.R. Morris. 2006. Assessment of public health risks associated with atmospheric exposure to PM 2.5 in Washington, DC, USA. *International Journal of Environmental Research and Public Health* 3(1): 86-97.

Griffin S., P. Goodrum, G. Diamond, W. Meylan, W. Brattin, and J. Hassett. 1999. Application of a probabilistic risk assessment methodology to a lead smelter site. *Human and Ecological Risk Assessment* 5(4): 845-868.

Griffitt R.J., J. Luo, J. Gao, J.C. Bonzongo, and D.S. Barber. 2008. Effects of particle composition and species on toxicity of metallic nanomaterials in aquatic organisms. *Environmental Toxicology and Chemistry* 27(9): 1972-1978.

Grist E.P.M., M. Crane, C. Jones, and P. Whitehouse. 2003. Estimation of demographic toxicity through the double bootstrap. *Water Research* 37(3): 618-626.

Gurjar B.R. and M. Mohan. 2003. Potential health risks due to toxic contamination in the ambient environment of certain Indian states. *Environmental Monitoring and Assessment* 82(2): 203-223.

Guy M., L. Singh, and P. Mineau. 2011. Using field data to assess the effects of pesticides on crustacea in freshwater aquatic ecosystems and verifying the level of protection provided by water quality guidelines. *Integrated Environmental Assessment and Management* 7(3): 426-436.

Hacon S., E.R.R. Rochedo, R.R.R. Campos, and L.D. Lacerda. 1997. Mercury exposure through fish consumption in the urban area of Alta Floresta in the Amazon Basin. *Journal of Geochemical Exploration* 58(2-3): 209-216.

Hall L.W. and J.M. Giddings. 2000. The Need for Multiple Lines of Evidence for Predicting Site-Specific Ecological Effects. *Human and Ecological Risk Assessment: An International Journal* 6(4): 679-710.

Hamilton A.J., F. Stagnitti, R. Premier, A.M. Boland, and G. Hale. 2006. Quantitative microbial risk assessment models for consumption of raw vegetables irrigated with reclaimed water. *Applied and Environmental Microbiology* 72(5): 3284-3290.

Handmer J., T. Norton, and S. Dovers. 2001.*Ecology, Uncertainty and Policy: managing ecosystems for sustainability*. Harlow: Prentice Hall.

Handy R.D., T.B. Henry, T.M. Scown, B.D. Johnston, and C.R. Tyler. 2008. Manufactured nanoparticles: their uptake and effects on fish - a mechanistic analysis. *Ecotoxicology* 17: 396-409.

Hansen Jesse L.C. and J.J. Obrycki. 2000. Field deposition of Bt transgenic corn pollen: Lethal effects on the monarch butterfly. *Oecologia* 125(2): 241-248.

Hart A., W. Roelofs, J. Crocker, A. Murray, N. Boatman, S. Hugo, S. Fitzpatrick, and V. Flari. *DEFRA Research Contract CPEC38: Quantitative approaches to the Risk Assessment of GM crops*. York (UK): Central Science Laboratory.

Hartigan J.A. 1975. *Clustering algorithms*. New York: John Wiley & Sons, Inc.

Hayes K., H. Regan, M. Burgman, R. Devlin, A. Kapuscinski, and J. Dambacher. 2006. Methods to address uncertainty in ecological risk assessment of Genetically Modified Organisms, *in* 'Ninth International Symposium on Biosafety of Genetically Modified

Organisms Biosafety Research and Environmental Risk Assessment', September 24-29, Jeju Island, South Korea.

Hayes E.H. and W.G. Landis. 2004. Regional ecological risk assessment of a near shore marine environment: Cherry Point, WA. *Human and Ecological Risk Assessment* 10(2): 299-325.

Hays S.M., L.L. Aylward, M. Gagné, and K. Krishnan. 2009. Derivation of Biomonitoring Equivalents for cyfluthrin. *Regulatory Toxicology and Pharmacology* 55(3): 268-275.

Health and Safety Executive. 2001. *Reducing risks, protecting people: HSE's decision-making process*. London: HMSO.

Hellmich R.L., B.D. Siegfried, M.K. Sears, D.E. Stanley-Horn, M.J. Daniels, H.R. Mattila, T. Spencer, K.G. Bidne, and L.C. Lewis. 2001. Monarch larvae sensitivity to Bacillus thuringiensis-purified proteins and pollen. *Proceedings of the National Academy of Sciences of the United States of America* 98(21): 11925-11930.

Helton J.C. and F.J. Davis. 2003. Latin Hypercube sampling and the propagation of uncertainty analysis of complex systems. *Reality Engineering and System Safety* 81: 23-69.

Helton J.C. 1994. Treatment of uncertainty in performance assessments for complex systems. *Risk Analysis* 14(4): 483-511.

Henning-De Jong I., R. Van Zelm, M.A.J. Huijbregts, D. De Zwart, T.M.A. van der Linden, A. Wintersen, L. Posthuma, and D. van de Meent. 2008. Ranking of agricultural pesticides in the Rhine-Meuse-Scheldt basin based on toxic pressure in marine ecosystems. *Environmental Toxicology and Chemistry* 27(3): 737-745.

Henrion M. and B. Fischhoff, B. 1986. Assessing uncertainty in physical constants. *American Journal of Physics* 54: 791.

Hoffman F. and J. Hammonds. 1994. Propagation of uncertainty in risk assessments: the need to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability. *Risk Analysis* 14(5): 707-712.

Hristozov D.R., S. Gottardo, A. Critto, and A. Marcomini. 2012. Risk assessment of engineered nanomaterials: a review of available data and approaches from a regulatory perspective. *Nanotoxicology* [epub ahead of print].

Hromkovic J. 2005. *Design and analysis of randomized algorithms: introduction to design paradigms*. New York: Springer-Verlag New York Inc.

Hughes K., M.E. Meek, M. Walker, and R. Beauchamp. 2003. 1,3-Butadiene: Exposure estimation, hazard characterization, and exposure-response analysis. *Journal of Toxicology and Environmental Health - Part B: Critical Reviews* 6(1): 55-83.

Huijbregts M., G. Norris, R. Bretz, A. Ciroth, B. Maurice, B. von Bahr, B. Weidema, and A. de Beaufort. 2001. Framework for modelling data uncertainty in life cycle inventories. *The International Journal of Life Cycle Assessment* 6(3): 127-132.

Hung M.L., S.Y. Wu, Y.C. Chen, H.C. Shih, Y.H. Yu, and H.W. Ma. 2009. The Health Risk Assessment of Pb and Cr leachated from fly ash monolith landfill. *Journal of Hazardous Materials* 172(1): 316-323.

Huysmans M., T. Madarász, and A. Dassargues. 2006. Risk assessment of groundwater pollution using sensitivity analysis and a worst-case scenario analysis. *Environmental Geology* 502: 180-193.

Jackson L.E., S.L. Bird, R.W. Matheny, R.V. O'Neill, D. White, K.C. Boesch, and J.L. Koviach. 2004. A regional approach to projecting land-use change and resulting ecological vulnerability. *Environmental Monitoring and Assessment* 94(1-3): 231-248.

Jackson J.D. 1996. Analysing the New Evidence Scholarship: Towards a New Conception of the Law of Evidence. *Oxford J Legal Studies* 16(2): 309-328.

Jahn H.J., A. Schneider, S. Breitner, R. Eißer, M. Wendisch, and A. Kräer. 2011. Particulate Matter Pollution in the Megacities of the Pearl River Delta, China – A Systematic Literature Review and Health Risk Assessment. *International Journal of Hygiene and Environmental Health* 214: 281-295.

Janssen P.H.M., A.C. Petersen, J.P. van der Sluijs, J.S. Risbey, and J.R. Ravetz. 2003. RIVM/MNP guidance for uncertainty assessment and communication: Quickscan hints and actions list. *RIVM/MNP Guidance for Uncertainty Assessment and Communication Series, volume 2*.

Jensen F.V. 1996. *An Introduction to Bayesian Networks*. London: University College London Press.

Jiménez E., C. Linares, L.F. Rodríguez, M.J. Bleda, and J. Díaz. 2009. Short-term impact of particulate matter (PM2.5) on daily mortality among the over-75 age group in Madrid (Spain). *Science of the Total Environment* 407 (21): 5486-5492.

Johnson A.C., M.J. Bowes, A. Crossley, H.P. Jarvie, K. Jurkschat, M.D. Jürgens, A.J. Lawlor, B. Park, P. Rowland, and D. Spurgeon. 2011. An assessment of the fate, behaviour and environmental risk associated with sunscreen TiO2 nanoparticles in UK field scenarios. *Science of the Total Environment* 409:(13) 2503-2510.

Jones D.R., J.L. Peters, L. Rushton, A.J. Sutton, and K.R. Abrams. 2009. Interspecies extrapolation in environmental exposure standard setting: A Bayesian synthesis approach. *Regulatory Toxicology and Pharmacology* 53(3): 217-225.

Jorgensen R., T. Hauser, T. D'Hertefeldt, N. Andersen, and D. Hooftman. 2009. The variability of processes involved in transgene dispersal - case studies from Brassica and related genera. *Environmental Science and Pollution Research* 16(4): 389-395.

Kahneman D., P. Slovic, and A. Tversky. 1982. *Judgement under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.

Kaloudis S., A. Tocatlidou, N.A. Lorentzos, A.B. Sideridis, and M. Karteris. 2005. Assessing wildfire destruction danger: a decision support system incorporating uncertainty. *Ecological Modelling* 181(1): 25-38.

Kandlikar M., G. Ramachandran, A. Maynard, B. Murdock, and .A. Toscano. 2007. Health risk assessment for nanoparticles: A case for using expert judgment. *Nanotechnology and Occupational Health* 91: 137-156.

Kapo K.E. and G.A. Burton Jr. 2006. A geographic information systems-based, weights-of-evidence approach for diagnosing aquatic ecosystem impairment. *Environmental Toxicology and Chemistry* 25(8): 2237-2249.

Keiter S., T. Braunbeck, S. Heise, S. Pudenz, W. Manz, and H. Hollert. 2009. A fuzzy logic-classification of sediments based on data from in vitro biotests. *Journal of Soils and Sediments* 9(3): 168-179.

Keith, D. 1996. When is it appropriate to Combine Expert Judgements? An Editorial Essay**.** *Climate Change* 33: 139-143.

Kelly M., H. Bennion, A. Burgess, J. Ellis, S. Juggins, R. Guthrie, J. Jamieson, V. Adriaenssens, and M. Yallop. 2009. Uncertainty in ecological status assessments of lakes and rivers using diatoms. *Hydrobiologia* 633(1): 5-15.

Kelly E. and K. Campbell. 2000. Separating varaibility and uncertainty in environmental risk assessment – making choices. *Human and Ecological Risk Assessment* 6(1): 1-13.

Kentel E. and M.M. Aral. 2007. Risk tolerance measure for decision-making in fuzzy analysis: A health risk assessment perspective. *Stochastic Environmental Research and Risk Assessment* 21(4): 405-417.

van der Keur P., H. Henriksen, J. Refsgaard, M. Brugnach, C. Pahl-Wostl, A. Dewulf, and H. Buiteveld. 2008. Identification of major sources of uncertainty in current IWRM practice. Illustrated for the Rhine basin. *Water Resources Management* 22(11): 1677-1708.

Khan F., R. Sadiq, and T. Husain. 2002. GreenPro-I: a risk-based life cycle assessment and decision-making methodology for process plant design. *Environmental Modelling & Software* 17(8): 669-692.

King R.S. and C.J. Richardson. 2003. Integrating bioassessment and ecological risk assessment: an approach to developing numerical water-quality criteria. *Environmental Management* 31: 795-809.

Kirk G.S., J.E. Raven, and M. Schofield. 1962. *The Presocratic Philosophers: A Critical History With a Selection of Texts*. Cambridge (UK): Cambridge University Press.

Klaine S.J., P.J.J. Alvarez, G.E. Batley, T.F. Fernandes, R.D. Handy, D.Y. Lyon, S. Mahendra, M.J. McLaughlin, and J.R. Lead. 2008. Nanomaterials in the environment: behaviour, fate, bioavailability, and effects. *Environmental Toxicology and Chemistry* 27(9): 1825-1851.

Klier C., S. Grundmann, S. Gayler, and E. Priesack. 2008. Modelling the environmental fate of the herbicide glyphosate in soil lysimeters. *Water, Air, and Soil Pollution: Focus* 82: 187-207.

Knol A.B., P. Slottje, J.P. van der Sluijs, and E. Lebret. 2010. The use of expert elicitation in environmental health impact assessment: a seven step procedure. *Environmental Health* 9(19).

Knol A.B., A.C. Petersen, J.P. van der Sluijs, and E. Lebret. 2009a. Dealing with uncertainties in environmental burden of disease assessment. *Environmental Health: A Global Access Science Source* 8(21).

Knol A.B., J.J. de Hartog, H. Boogaard, P. Slottje, J.P. van der Sluijs, E. Lebret, F.R. Cassee, J.A. Wardekker, J.G. Ayres, P.J. Borm, *et al.* 2009b. Expert elicitation on ultrafine particles: likelihood of health effects and causal pathways. *Particle and Fibre Toxicology* 6(19).

Kooistra L., M.A.J. Huijbregts, A.M.J. Ragas, R. Wehrens, and R.S.E.W. Leuven. 2005. Spatial variability and uncertainty ecological risk assessment: A case study on the potential risk of cadmium for the little owl in a Dutch river flood plain. *Environmental Science and Technology* 39(7): 2177-2187.

Kotra J.P., M.P Lee, N.A. Eisenberg, and A.R. DeWispelare. 1996. *Branch technical position on the use of expert elicitation in the high-level radioactive waste program*. Washington: U.S. Nuclear Regulatory Commission.

Krayer von Krauss M. 2005. *Uncertainty in policy relevant sciences*. PhD thesis, Technical University of Denmark.

Krayer von Krauss M., E.A. Casman, and M.J. Small. 2004. Elicitation of expert judgments of uncertainty in the risk assessment of herbicide-tolerant oilseed crops. *Risk Analysis* 24:1515-1527.

Krimsky S. 2005. The weight of scientific evidence in policy and law. *American Journal of Public Health* 95 (Suppl. 1): S129–S136.

Kumar V., M. Mari, M. Schuhmacher, and J. Domingo. 2009. Partitioning total variance in risk assessment: Application to a municipal solid waste incinerator. *Environmental Modelling & Software* 24(2): 247-261.

Laden F., L.M. Neas, D.W. Dockery, and J. Schwartz. 2000. Association of fine particulate matter from different sources with daily mortality in six U.S. cities. *Environmental Health Perspectives* 108(10): 941-947.

Lai C.-H., S.-H. Liou, T.-S. Shih, P.-J. Tsai, H.-L. Chen, Y.-C. Chang, T.J. Buckley, P. Strickland, and J.J.K. Jaakkola. 2004. Exposure to fine particulate matter (PM2.5) among highway toll station workers in Taipei: Direct and indirect exposure assessment. *Archives of Environmental Health* 59(3): 138-148.

Landis W.G. 2005. Population-scale assessment endpoints in ecological risk assessment part 1: Reflections of stakeholder values. *Integrated Environmental Assessment and Management* 2(1): 86-91.

Landis W.G., P.B. Duncan, E.H. Hayes, A.J. Markiewicz, and J.F. Thomas. 2004. A regional retrospective assessment of the potential stressors causing the decline of the cherry point pacific herring run. *Human and Ecological Risk Assessment* 10(2): 271-297.

Lapresta-Fernández A., A. Fernández, and J. Blasco. 2012. Nanoecotoxicity effects of engineered silver and gold nanoparticles in aquatic organisms. *TrAC - Trends in Analytical Chemistry* 32: 40-59.

Lee J.J., C.S. Jang, C.P. Liang, and C.W. Liu. 2008. Assessing carcinogenic risks associated with ingesting arsenic in farmed smeltfish (Ayu, Plecoglossus altirelis) in aseniasis-endemic area of Taiwan. *Science of the Total Environment* 403 1-3): 68-79.

Lemke L.D. and A.S. Bahrou. 2009. Partitioned multiobjective risk modeling of carcinogenic compounds in groundwater. *Stochastic Environmental Research and Risk Assessment* 23(1): 27-39.

Li H.L., G.H. Huang, and Y. Zou. 2008. An integrated fuzzy-stochastic modeling approach for assessing health-impact risk from air pollution. *Stochastic Environmental Research and Risk Assessment* 22(6): 789-803.

Li J., G. Huang, G. Zeng, I. Maqsood, and Y. Huang. 2007. An integrated fuzzy-stochastic modeling approach for risk assessment of groundwater contamination. *Journal of Environmental Management* 82(2): 173-188.

Li J., L. Liu, G. Huang, and G. Zeng. 2006. A fuzzy-set approach for addressing uncertainties in risk assessment of hydrocarbon-contaminated site. *Water, Air, and Soil Pollution* 171(1-4): 5-18.

Liao C.M. and B.Y.H. Chou. 2005. Predictive risk thresholds for survival protection of farmed abalone, Haliotis diversicolor supertexta, exposed to waterborne zinc. *Environmental Toxicology* 20(2): 202-211.

Lindenschmidt K.E., S. Huang, and M. Baborowski. 2008. A quasi-2D flood modeling approach to simulate substance transport in polder systems for environment flood risk assessment. *Science of the Total Environment* 397(1-3): 86-102.

Lindley D. 2008. *Uncertainty: Einstein, Heisenberg, Bohr, and the Struggle for the Soul of Science*. New York: Anchor.

Linkov I., D. Loney, S. Cormier, F. Satterstrom, and T. Bridges. 2009. Weight-of-evidence evaluation in environmental assessment: Review of qualitative and quantitative approaches. *Science of the Total Environment* 407(19): 5199-5205.

Linkov I., F.K. Satterstrom, J. Steevens, E. Ferguson, and R.C. Pleus. 2007. Multi-criteria decision analysis and environmental risk assessment for nanomaterials. *Journal of Nanoparticle Research* 94: 543-554.

Linkov I. and D. Burmistrov. 2005. Sources of uncertainty in model predictions: Lessons learned from the IAEA Forest and Fruit Working Group model intercomparisons. *Journal of Environmental Radioactivity* 84(2): 297-314.

Linkov I., K.E. Von Stackelberg, D. Burmistrov, and T.S. Bridges. 2001. Uncertainty and variability in risk from trophic transfer of contaminants in dredged sediments. *Science of The Total Environment* 274(1-3): 255-269.

Liu Z.Q., Y.H. Zhang, G.H. Li, and X. Zhang. 2007. Sensitivity of key factors and uncertainties in health risk assessment of benzene pollutant. *Journal of Environmental Sciences* 19(10): 1272-1280.

Loos M., A.M.J. Ragas, R. Plasmeijer, A.M. Schipper, and A.J. Hendriks. 2010. Eco-SpaCE: An object-oriented, spatially explicit model to assess the risk of multiple environmental stressors on terrestrial vertebrate populations. *Science of the Total Environment* 408(18): 3908-3917.

Lorenz C., H. Hagendorfer, N. von Goetz, R. Kaegi, R. Gehrig, A. Ulrich, M. Scheringer, and K. Hungerbühler. 2011. Nanosized aerosols from consumer sprays: Experimental analysis and exposure modeling for four commercial products. *Journal of Nanoparticle Research* 13(8): 3377-3391.

Lorenz E.N. 1995. *The essence of chaos*. London: University College London.

Losey J.E., L.S. Rayor, and M.E. Carter. 1999. Transgenic pollen harms monarch larvae. *Nature* 399(6733): 214.

Loveridge D. 2004. Experts and foresight: review and experience. *International Journal of Foresight and Innovation Policy* 1: 33-69.

Lu H., L. Axe, T.A. Tyson. 2003. Development and application of computer simulation tools for ecological risk assessment. *Environmental Modeling and Assessment* 8(4): 311-322.

Ma H.W. 2002. Stochastic multimedia risk assessment for a site with contaminated groundwater. *Journal of Environmental Management* 166: 464-478.

Ma W.C. and H. van der Voet. 1993. A risk-assessment model for toxic exposure of small mammalian carnivores to cadmium in contaminated natural environments. *Science of the Total Environment* (Suppl. 2): 1701-1714.

Madl A.K. and K.E. Pinkerton. 2009. Health effects of inhaled engineered and incidental nanoparticles. *Critical Reviews in Toxicology* 39: 629-658.

Maier H., J. Ascough II, M. Wattenbach, C. Renschler, W. Labiosa, and J. Ravalico. 2008. Chapter Five - Uncertainty in Environmental Decision Making: Issues, Challenges and Future Directions. *Developments in Integrated Environmental Assessment* 3: 69-85.

Manikandan S. 2011. Measures of dispersion. *Journal of Pharmacology and Pharmacotherapeutics* 2: 315-316.

Martonen T.B. and J.D. Schroeter. 2003. Risk assessment dosimetry model for inhaled particulate matter: I. Human subjects. *Toxicology Letters* 138(1-2): 119-132.

Mastin B.J., and J.H. Rodgers Jr. 2000. Toxicity and bioavailability of copper herbicides (clearigate, cutrine-plus, and copper sulfate) to freshwater animals. *Archives of Environmental Contamination and Toxicology* 39(4): 445-451.

Matranga V.and I. Corsi. 2012. Toxic effects of engineered nanoparticles in the marine environment: Model organisms and molecular approaches. *Marine Environmental Research* 76: 32-40.

Matson C.W., A.M. Gillespie, C. McCarthy, T.J. McDonald, J.W. Bickham, R. Sullivan, and K.C. Donnelly. 2009. Wildlife toxicology: Biomarkers of genotoxic exposures at a hazardous waste site. *Ecotoxicology* 18(7): 886-898.

Mattila H.R., M.K. Sears, and J.J. Duan. 2005. Response of Danaus plexippus to pollen of two new Bt corn events via laboratory bioassay. *Entomologia Experimentalis et Applicata* 116(1): 31-41.

Maxim L.D. and E.E. McConnell. 2001. Interspecies comparisons of the toxicity of asbestos and synthetic vitreous fibers: A weight-of-the-evidence approach. *Regulatory Toxicology and Pharmacology* 33(3): 319-342.

Maxwell R.M. and W.E. Kastenberg. 1999. Stochastic environmental risk analysis: An integrated methodology for predicting cancer risk from contaminated groundwater. *Stochastic Environmental Research and Risk Assessment* 13(1-2): 27-47.

Maycock B.J. and D.J. Benford. 2007. Risk assessment of dietary exposure to methylmercury in fish in the UK. *Human and Experimental Toxicology* 26(3): 185-190.

McColl S., J. Hicks, L. Craig, and J. Shortreed. 2000. *Environmental Health Risk Management: A Primer for Canadians*. University of Waterloo, Institute for Risk Research.

McDonald B.G. and J.B. Wilcockson. 2003. Improving the use of toxicity reference values in wildlife food chain modeling and ecological risk assessment. *Human and Ecological Risk Assessment* 9(7): 1585-1594.

Meek M.E., R. Beauchamp, G. Long, D. Moir, L. Turner, and M. Walker. 2002. Chloroform: Exposure estimation, hazard characterization, and exposure-response analysis. *Journal of Toxicology and Environmental Health - Part B: Critical Reviews* 5(3): 283-334.

Meek M.E. and K. Hughes. 1995. Approach to health risk determination for metals and their compounds under the Canadian Environmental Protection Act. *Regulatory Toxicology and Pharmacology* 22(3): 206-212.

Menzie C., M.H. Henning, J. Cura, K. Finkelstein, J. Gentile, J. Maughan, D. Mitchell, S. Petron, B. Potocki, S. Svirsky, and P. Tyler. 1996. Special report of the Massachusetts weight-of-evidence workgroup: A weight-of-evidence approach for evaluating ecological risks. *Human and Ecological Risk Assessment* 2(2): 277-304.

Merz B. and A.H. Thieken. 2009. Flood risk curves and uncertainty bounds. *Natural Hazards* 513: 437-458.

Meyer V., S. Scheuer, and D. Haase. 2009. A multicriteria approach for flood risk mapping exemplified at the Mulde river, Germany. *Natural Hazards* 48(1): 17-39.

Meyer M. and J. Booker. 1991. *Eliciting and Analyzing Expert Judgment: A Practical Guide*. London: Academic Press.

Morgan M., M. Henrion, and M. Small. 1990. *Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis*. Cambridge: Cambridge University Press.

Mueller N.C. and B. Nowack. 2008. Exposure modeling of engineered nanoparticles in the environment. *Environmental Science and Technology* 42: 4447-4453.

Mugglestone M.A., E.D. Stutt, and L. Rushton. 2001. Setting microbiological water quality standards for sea bathing - A critical evaluation. *Water Science and Technology* 43(12): 9-18.

Mukhtasor H.T., B. Veitch, and N. Bose. 2004. An ecological risk assessment methodology for screening discharge alternatives of produced water. *Human and Ecological Risk Assessment* 10(3): 505-524.

Musee N. 2011. Simulated environmental risk estimation of engineered nanomaterials: A case of cosmetics in Johannesburg City. *Human and Experimental Toxicology* 30(9): 1181-1195.

Musee N., P.J. Oberholster, L. Sikhwivhilu, and A.M. Botha. 2010. The effects of engineered nanoparticles on survival, reproduction, and behaviour of freshwater snail, Physa acuta (Draparnaud, 1805). *Chemosphere* 81: 1196-1203.

Naito W., Y. Gamo, K. Yoshida. 2006. Screening-level risk assessment of Di(2-ethylhexyl)phthalate for aquatic organisms using monitoring data in Japan. *Environmental Monitoring and Assessment* 115(1-3): 451-471.

Nasiri F., G. Huang, and N. Fuller. 2007. Prioritizing groundwater remediation policies: A fuzzy compatibility analysis decision aid. *Journal of Environmental Management* 82(1): 13-23.

National Research Council/National Academy of Sciences. 1983. *Risk Assessment in the Federal Government: Managing the Process*. Washington: National Academy Press.

de Nazelle A., and D.A. Rodríguez. 2009. Tradeoffs in incremental changes towards pedestrian-friendly environments: Physical activity and pollution exposure. *Transportation Research Part D: Transport and Environment* 14(4): 255-263.

Nazir M., F. Khan, P. Amyotte, and R. Sadiq. 2008. Subsea release of oil from a riser: an ecological risk assessment. *Risk Analysis* 28(5): 1173-1196.

Neuhäuser B. and B. Terhorst. 2007. Landslide susceptibility assessment using "weights-of-evidence" applied to a study area at the Jurassic escarpment (SW-Germany). *Geomorphology* 86(1-2): 12-24.

de Nijs T.C.M., C. Toet, T.G. Vermeire, P. van der Poel, and J. Tuinstra. 1993. Dutch risk assessment system for new chemicals: DRANC. *Science of the Total Environment* (Suppl. 2): 1729-1748.

Oberhauser K.S., M.D. Prysby, H.R. Mattila, D.E. Stanley-Horn, M.K. Sears, G. Dively, E. Olson, J.M. Pleasants, W.-K.F. Lam, and R.L. Hellmich. 2001. Temporal and spatial overlap between monarch larvae and corn pollen. *Proceedings of the National Academy of Sciences of the United States of America* 98(21): 11913-11918.

Olson M.S. and P.L. Gurian. 2012. Risk assessment strategies as nanomaterials transition into commercial applications. *Journal of Nanoparticle Research* 14(4).

van der Oost R., J. Beyer, and N.P.E. Vermeulen. 2003. Fish bioaccumulation and biomarkers in environmental risk assessment: A review. *Environmental Toxicology and Pharmacology* 13(2): 57-149.

Orru H., M. Maasikmets, T. Lai, T. Tamm, M. Kaasik, V. Kimmel, K. Orru, E. Merisalu, and B. Forsberg. 2011. Health impacts of particulate matter in five major Estonian towns:

Main sources of exposure and local differences. *Air Quality, Atmosphere and Health* 4(3): 247-258.

Oughton D.H., A. Agüero, R. Avila, J.E. Brown, D. Copplestone, M. Gilek. 2008. Addressing uncertainties in the ERICA Integrated Approach. *Journal of Environmental Radioactivity* 99: 1384-1392.

Palma G., A. Sanchez, Y. Olave, F. Encina, R. Palma, and R. Barra. 2004. Pesticide levels in surface waters in an agricultural-forestry basin in Southern Chile. *Chemosphere* 57(8): 763-770.

Park R.A., J.S. Clough, and M.C. Wellman. 2008. AQUATOX: Modeling environmental fate and ecological effects in aquatic ecosystems. *Ecological Modelling* 213(1): 1-15.

Pascoe G.A., R.J. Blanchet, and G. Linder. 1993. Ecological risk assessment of a metals-contaminated wetland: Reducing uncertainty. *Science of the Total Environment* (Suppl. 2): 1715-1728.

Paustenbach D.J. 1989. A survey of health risk assessment, *in* Paustenbach D.J. (ed) *The risk assessment of environmental and human health hazards: a textbook of case studies*. New York: John Wiley and Son.

Payne S. 1951. *The Art of Asking Questions*. Princeton: Princeton University Press.

Perrot P. 1998. *A to Z of Thermodynamics*. Oxford: Oxford University Press.

Perry J.N., Y. Devos, S. Arpaia, D. Bartsch, A. Gathmann, R.S. Hails, J. Kiss, K. Lheureux, B. Manachini, S. Mestdagh, *et al*. 2010. A mathematical model of exposure of nontarget Lepidoptera to Bt-maize pollen expressing Cry1Ab within Europe. *Proceedings of the Royal Society B: Biological Sciences* 277(1686): 1417-1425.

Persson K. and G. Destouni. 2009. Propagation of water pollution uncertainty and risk from the subsurface to the surface water system of a catchment. *Journal of Hydrology* 377(3-4): 434-444.

Petersen A.C. 2006. *Simulating nature: a philosophical study of computer-simulation uncertainties and their role in climate science and policy advice*. Amsterdam: Aksant Academic Publishers.

Peterson R.K.D., S.J. Meyer, A.T. Wolf, J.D. Wolt, and P.M. Davis. 2006. Genetically engineered plants, endangered species and risk: a temporal and spatial exposure assessment for Karner blue butterfly larvae and Bt maize pollen. *Risk Analysis* 26: 845-858.

Pollard S.J.T. 2001. An overview of the use of risk assessment for environmental regulation in the UK - key drivers and regulatory initiatives. *Risk Decision and Policy* 6(1): 33-46.

Pollino C.A., O. Woodberry, A. Nicholson, K. Korb, and B.T. Hart. 2007. Parameterisation and evaluation of a Bayesian network for use in an ecological risk assessment. *Environmental Modelling and Software* 22(8): 1140-1152.

Pope III C.A. and D.W. Dockery. 2006. Health effects of fine particulate air pollution: lines that connect. *Journal of the Air and Waste Management Association* 56: 709-42.

Post E., D. Hoaglin, L. Deck, and K. Larntz. 2001. An empirical bayes approach to estimating the relation of mortality to exposure to particulate matter. *Risk Analysis* 21(5): 837-842.

Phillips K.P., W.G. Foster, W. Leiss, V. Sahni, N. Karyakina, M.C. Turner, S. Kacew, and D. Krewski. 2008. Assessing and managing risks arising from exposure to endocrine-active chemicals. *Journal of Toxicology and Environmental Health, Part B: Critical Reviews* 113-4: 351-372.

Proctor D.M., E.C. Shay, K.A. Fehling, and B.L. Finley. 2002. Assessment of human health and ecological risks posed by the uses of steel-industry slags in the environment. *Human and Ecological Risk Assessment* 8(4): 681-711.

Prudhomme C., D. Jakob, and C. Svensson. 2003. Uncertainty and climate change impact on the flood regime of small UK catchments. *Journal of Hydrology* 277(1-2): 1-23.

Qin X. and G. Huang. 2009. Characterizing uncertainties associated with contaminant transport modeling through a coupled fuzzy-stochastic approach. *Water, Air, & Soil Pollution* 197(1): 331-348.

Quik J.T.K., J.A. Vonk, S.F. Hansen, A. Baun, and D. van de Meent. 2011. How to assess exposure of aquatic organisms to manufactured nanoparticles? *Environment International* 37(6): 1068-1067.

Ranke J. 2002. Persistence of antifouling agents in the marine biosphere. *Environmental Science and Technology* 36(7): 1539-1545.

Rauschen S., F. Schaarschmidt, and A. Gathmann. 2010. Occurrence and field densities of Coleoptera in the maize herb layer: implications for environmental risk assessment of genetically modified Bt-maize. *Transgenic Research* 19: 727-744.

Ravnum S., K.E. Zimmer, H. Keune, A.C. Gutleb, A.J. Murk, J.G. Koppe, B.L. Magnanti, J.L. Lyche, G.S. Eriksen E. Ropstad, *et al.* 2012. Policy relevant Results from an Expert Elicitation on the Human Health Risks of Decabromo-diphenyl ether (decaBDE) and Hexabromocyclododecane (HBCD). *Environmental Health* 11(Suppl. 1): S7.

Raybould A. Problem formulation and hypothesis testing for environmental risk assessments of genetically modified crops. *Environmental Biosafety Research.* 5: 119-125.

Refsgaard J.C., J.P. van der Sluijs, A. Hojberg, and P. Vanrolleghem. 2007. Uncertainty in the environmental modelling process - A framework and guidance. *Environmental Modelling & Software* 22(11): 1543-1556.

Regan H., M. Colyvan, and M. Burgman. 2002. A taxonomy and treatment of uncertainty for ecology and conservation biology. *Ecological Applications* 12(2): 618-628.

Renn O. and M.C. Roco. 2006. Nanotechnology and the need for risk governance. *Journal of Nanoparticle Research* 8(2): 153-191.

Rocks S.A., R.A. Dorey, P.T.C. Harrison, L.S. Levy, R.D. Handy, and S.J.T. Pollard. 2008. *Comparison of risk assessment approaches for manufactured nanomaterials*. Client report for DEFRA (Project CB403).

Rowe G. and G. Wright. 1999. The Delphi technique as a forecasting tool: issues and analysis. *International Journal of Forecasting* 15(4): 353-375.

Rowe W. 1994. Understanding uncertainty. *Risk Analysis* 14(5): 743-750.

Royal Society. 1992. *Risk Analysis, Perception and Management*. London: Royal Society.

Rutherford L., C. Garron, W. Ernst, and K. Kennedy. 2003. The aquatic environment and textile mill effluents - An ecological risk assessment. *Human and Ecological Risk Assessment* 9(2): 589-606.

Saldarriaga-Noreña H., L. Hernández-Mena, M. Ramírez-Muñiz, P. Carbajal-Romero, R. Cosío-Ramírez, and B. Esquivel-Hernández. 2009. Characterization of trace metals of risk to human health in airborne particulate matter (PM2.5) at two sites in Guadalajara, Mexico. *Journal of Environmental Monitoring* 11(4): 887-894.

Sander P. and T. Öberg. 2006. Comparing deterministic and probabilistic risk assessments: A case study at a closed steel mill in southern Sweden. *Journal of Soils and Sediments* 6(1): 55-61.

Sanderson H., B. Laird, L. Pope, R. Brain, C. Wilson, D. Johnson, G. Bryning, A.S. Peregrine, A. Boxall, and K. Solomon. 2007. Assessment of the environmental fate and effects of ivermectin in aquatic mesocosms. *Aquatic Toxicology* 85(4): 229-240.

Sanderson H., S.D. Dyer, B.B. Price, A.M. Nielsen, R. van Compernolle, M. Selby, K. Stanton, A. Evans, M. Ciarlo, and R. Sedlak. 2006. Occurrence and weight-of-evidence risk assessment of alkyl sulfates, alkyl ethoxysulfates, and linear alkylbenzene sulfonates (LAS) in river water and sediments. *Science of the Total Environment* 368(2-3): 695-712.

Scherm H. 2000. Simulating uncertainty in climate-pest models with fuzzy numbers. *Environmental Pollution* 108(3): 373-379.

Schoeny R., L. Haber, and M. Dourson. 2006. Data considerations for regulation of water contaminants. *Toxicology* 221(2-3): 217-224.

Schuler L.J. and G.M. Rand. 2008. Aquatic risk assessment of herbicides in freshwater ecosystems of south Florida. *Archives of Environmental Contamination and Toxicology* 54(4): 571-583.

Schwartz J. 2002. The use of epidemiology in environmental risk assessment. *Human and Ecological Risk Assessment* 8: 1253-1265.

Schwartz S., V. Berding, and M. Matthies. 2000. Aquatic fate assessment of the polycyclic musk fragrance HHCB scenario and variability analysis in accordance with the EU risk assessment guidelines. *Chemosphere* 41(5): 671-679.

Scientific Committee on Emerging and Newly Identified Health Risks. 2009. *Risk Assessment of Products of Nanotechnologies*. Brussels: European Commission.

Scott M.J., C.A. Brandt, A.L. Bunn, D.W. Engel, P.W. Eslinger, T.B. Miley, B.A. Napier, L.A. Nieves, and E.L. Prendergast. 2005. Modeling long-term risk to environmental and human systems at the Hanford nuclear reservation: Scope and findings from the initial model. *Environmental Management* 35(1): 84-98.

Sears M.K., R.L. Hellmich, D.E. Stanley-Horn, K.S. Oberhauser, J.M. Pleasants, H.R. Mattila, B.D. Siegfried, and G.P. Dively. 2001. Impact of Bt corn pollen on monarch butterfly populations: A risk assessment. *Proceedings of the National Academy of Sciences of the United States of America* 98(21): 11937-11942.

Shakhawat C., H. Tahir, and B. Neil. 2006. Fuzzy rule-based modelling for human health risk from naturally occurring radioactive materials in produced water. *Journal of Environmental Radioactivity* 89(1): 1-17.

Shaw B.J. and R.D. Handy. Physiological effects of nanoparticles on fish: A comparison of nanometals versus metal ions. *Environment International* 37(6): 1083-1097.

Siemering G.S., J.D. Hayworth, and B.K. Greenfield. 2008. Assessment of potential aquatic herbicide impacts to California aquatic ecosystems. *Archives of Environmental Contamination and Toxicology* 55(3): 415-431.

Slottje P., J.P. van der Sluijs, and A.B. Knol. 2008. Expert elicitation: *Methodological suggestions for its use in environmental health impacts assessments*. RIVM letter report 630004001/2008. Bilthoven, Netherlands: RIVM, Copernicus Institute for Sustainable Development and Innovation.

van der Sluijs J.P., M. Craye, S. Funtowicz, P. Kloprogge, J. Ravetz, and J.S. Risbey. 2005. Combining quantitative and qualitative measures of uncertainty in model-based environmental assessment: The NUSAP system. *Risk Analysis* 25(2): 481-492.

van der Sluijs J.P., P.H.M Janssen, A.C. Petersen, P. Kloprogge, J.S. Risbey, W. Tuinstra, and J.R. Ravetz. 2004. RIVM/MNP guidance for uncertainty assessment and communication: Tool catalogue for uncertainty assessment. *RIVM/MNP Guidance for Uncertainty Assessment and Communication Series, volume 2*.

van der Sluijs J.P. 1997. *Anchoring amid uncertainty, On the management of uncertainties in risk assessment of anthropogenic climate change*. PhD thesis, Utrecht University.

Smita S., S.K. Gupta, A. Bartonova, M. Dusinska, A.C. Gutleb, and Q. Rahman. 2012. Nanoparticles in the environment: assessment using the causal diagram approach. *Environmental Health* 11(Suppl. 1)**:** S13.

Smith J.T., L.A. Walker, R.F. Shore, S.E.A. Le v. Dit Durell, P.D. Howe, and M. Taylor. 2009. Do estuaries pose a toxic contamination risk for wading birds? *Ecotoxicology* 18(7): 906-917.

Smith P.N., G.P. Cobb, C. Godard-Codding, D. Hoff, S.T. McMurry, T.R. Rainwater, and K.D. Reynolds. 2007. Contaminant exposure in terrestrial vertebrates. *Environmental Pollution* 150(1): 41-64.

Society for Environmental Toxicology and Chemistry. 1994. *Aquatic Dialogue Group: Pesticide Risk Assessment and Mitigation.* Pensacola, FL: Society for Environmental Toxicology and Chemistry and SETAC Foundation for Environmental Education.

Som C., P. Wick, H. Krug, and B. Nowack. 2011. Environmental and health effects of nanomaterials in nanotextiles and façade coatings. *Environment International* 37(6): 1131-1142.

Son H.V., S. Ishihara, and H. Watanabe. 2006. Exposure risk assessment and evaluation of the best management practice for controlling pesticide runoff from paddy fields. Part 1: Paddy watershed monitoring. *Pest Management Science* 62(12): 1193-1206.

Speirs-Bridge A., F. Fidler, M. McBride, L. Flander, G. Cumming, and M.A. Burgman. 2010. Reducing overconfidence in the interval judgments of experts. *Risk Analysis* 30: 512-523.

van Sprang P.A., F.A.M. Verdonck, F. van Assche, L. Regoli, K.A.C. de Schamphelaere. 2009. Environmental risk assessment of zinc in European freshwaters: A critical appraisal. *Science of the Total Environment* 407(20): 5373-5391.

Stanley-Horn D.E., G.P. Dively, R.L. Hellmich, H.R. Mattila, M.K. Sears, R. Rose, L.C.H. Jesse, J.E. Losey, J.J. Obrycki, and L. Lewis. 2001. Assessing the impact of Cry1Ab-expressing corn pollen on monarch butterfly larvae in field studies. *Proceedings of the National Academy of Sciences of the United States of America* 98(21): 11931-11936.

Staples C.A., K. Woodburn, N. Caspers, A.T. Hall, and G.M. Klečka. 2002. A weight of evidence approach to the aquatic hazard assessment of bisphenol A. *Human and Ecological Risk Assessment* 8(5): 1083-1105.

Stephens M., B. Goodwin, and T. Andres. 1993. Deriving parameter probability density functions. *Reliability Engineering & System Safety* 42(2-3): 271-291.

Stevens G., B. de Foy, J.J. West, and J.I. Levy. 2007. Developing intake fraction estimates with limited data: Comparison of methods in Mexico City. *Atmospheric Environment* 41(17): 3672-3683.

Stirling A. 1998. Risk at a turning point? *Journal of Risk Research* 1(2): 97-109.

Sullivan J., N. Ishikawa, L. Sheppard, D. Siscovick, H. Checkoway, and J. Kaufman. 2003. Exposure to ambient fine particulate matter and primary cardiac arrest among persons with and without clinically recognized heart disease. *American Journal of Epidemiology* 157(6): 501-509.

Suter G.W. II. 1996. *Guide for developing conceptual models for ecological risk assessment*. Oak Ridge (Tennessee): U.S. Department of Energy, National Laboratory environmental restoration risk assessment program.

Symons J.M., L. Wang, E. Guallar, E. Howell, F. Dominici, M. Schwab, B.A. Ange, J. Samet, J. Ondov, D. Harrison, and A. Geyh. 2006. A case-crossover study of fine particulate matter air pollution and onset of congestive heart failure symptom exacerbation leading to hospitalization. *American Journal of Epidemiology* 164(5): 421-433.

Tainio M., J.T. Tuomisto, J. Pekkanen, N. Karvosenoja, K. Kupiainen, P. Porvari, M. Sofiev, A. Karppinen, L. Kangas, and J. Kukkonen. 2010. Uncertainty in health risks due to anthropogenic primary fine particulate matter from different source types in Finland. *Atmospheric Environment* 44(17): 2125-2132.

Teunis P.F.M., G.J. Medema, L. Kruidenier, and A.H. Havelaar. 1997. Assessment of the risk of infection by Cryptosporidium or Giardia in drinking water from a surface water source. *Water Research* 31(6): 1333-1346.

Therriault T.W. and L.M. Herborg. 2008. A qualitative biological risk assessment for vase tunicate Ciona intestinalis in Canadian waters: Using expert knowledge. *ICES Journal of Marine Science* 65(5): 781-787.

Thomas K.V., J. Farkas, E. Farmen, P. Christian, K. Langford, Q. Wu, and K.E. Tollefsen. 2011. Effects of dispersed aggregates of carbon and titanium dioxide engineered nanoparticles on rainbow trout hepatocytes. *Journal of Toxicology and Environmental Health - Part A: Current Issues* 74(7-9): 466-477.

Thomas T., T. Bahadori, N. Savage, and K. Thomas. 2009. Moving Towards Exposure and Risk Evaluation of Nanomaterials: Challenges and Future Directions. *Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology* 1: 426-433.

Thorsen M., J.C. Refsgaard, S. Hansen, E. Pebesma, J.B. Jensen, and S. Kleeschulte. 2001. Assessment of uncertainty in simulation of nitrate leaching to aquifers at catchment scale. *Journal of Hydrology* 242(3-4): 210-227.

Tillman Jr. F.D., and J.W. Weaver. 2006. Uncertainty from synergistic effects of multiple parameters in the Johnson and Ettinger (1991) vapor intrusion model. *Atmospheric Environment* 40(22): 4098-4112.

Thorsen M., J.C. Refsgaard, S. Hansen, E. Pebesma, J.B. Jensen, and S. Kleeschulte. 2001. Assessment of uncertainty in simulation of nitrate leaching to aquifers at catchment scale. *Journal of Hydrology* 242(3-4): 210-227.

Troldborg M. 2010. *Risk assessment models and uncertainty estimation of groundwater contamination from point sources*. PhD Thesis, Technical University of Denmark.

Tsuji J.S., R. Benson, R.A. Schoof, and G.C. Hook. 2004. Health effect levels for risk assessment of childhood exposure to arsenic. *Regulatory Toxicology and Pharmacology* 39(2): 99-110.

Twining J.R. and R.F. Cameron. 1997. Decision-making processes in ecological risk assessment using copper pollution of macquarie harbour from Mt. Lyell, Tasmania, as a case study. *Hydrobiologia* 352(1-3): 207-218.

United Nations. 1992. Rio declaration on environment and development, *in* 'Report of the United Nations conference on environment and development', June 3-14, Rio de Janeiro, Brazil.

United States Environmental Protection Agency. 2010. *Quantitative Health Risk Assessment for Particulate Matter*. Washington: United States Environmental Protection Agency Office of Air Quality Planning and Standards.

United States Environmental Protection Agency. 2009. *Expert elicitation task force white paper*. Washington: Science Policy Council.

United States Environmental Protection Agency. 2003. *Framework for cumulative risk assessment*. Washington: United States Environmental Protection Agency Risk Assessment Forum.

United States Environmental Protection Agency. 2002. *A Review of the Reference Dose and Reference Concentration Processes, EPA/630/P-02/002F*. Washington: United States Environmental Protection Agency Risk Assessment Forum.

United States Environmental Protection Agency. 1999. *ECOFRAM Terrestrial Draft Report*. Washington: Ecological Committee on FIFRA Risk Assessment Methods.

United States Environmental Protection Agency. 1998. *Guidelines for ecological risk assessment*. Washington: United States Environmental Protection Agency Risk Assessment Forum.

United States Environmental Protection Agency. 1997. *Exposure Factors Handbook*. Washington: United States Environmental Protection Agency Office of Research and Development.

United States Environmental Protection Agency. 1996. *Summary Report for the Workshop on Monte Carlo Analysis, EPA/630/R-96/010*. Washington: United States Environmental Protection Agency Risk Assessment Forum.

United States Environmental Protection Agency. 1992. *Framework for ecological risk assessment*. Washington: United States Environmental Protection Agency Risk Assessment Forum.

Uricchio V.F., R. Giordano, and N. Lopez. 2004. A fuzzy knowledge-based decision support system for ground water pollution risk evaluation. *Journal of Environmental Management* 73: 189-197.

Vallack H.W., D.J. Bakker, I. Brandt, E. Broström-Lundén, A. Brouwer, K.R. Bull, C. Gough, R. Guardans, I. Holoubek, B. Jansson, *et al.* 1998. Controlling persistent organic pollutants-what next? *Environmental Toxicology and Pharmacology* 6(3): 143-175.

Verdonck F.A.M., P.A. van Sprang, and P.A. Vanrolleghem. 2008. An intelligent data collection tool for chemical safety/risk assessment. *Chemosphere* 70(10): 1818-1826.

Vesely W. and D. Rasmuson. 1984. Uncertainties in nuclear probabilistic risk analyses. *Risk Analysis* 4(4): 313-322.

Vose D. 2000. *Risk Analysis*. Chichester (UK): John Wiley & Sons.

Vryzas Z., C. Alexoudis, G. Vassiliou, K. Galanis, and E. Papadopoulou-Mourkidou. 2011. Determination and aquatic risk assessment of pesticide residues in riparian drainage canals in northeastern Greece. *Ecotoxicology and Environmental Safety* 74(2): 174-181.

Vryzas Z., G. Vassiliou, C. Alexoudis, and E. Papadopoulou-Mourkidou. 2009. Spatial and temporal distribution of pesticide residues in surface waters in northeastern Greece. *Water Research* 43(1): 1-10.

Walker W., P. Harremoes, J. Rotmans, J. van der Sluijs, M.van Asselt, P. Janssen, and M. Krayer von Krauss. 2003. Defining uncertainty: a conceptual basis for uncertainty management in model-based decision support. *Integrated Assessment* 4(1): 5-17.

Walker R., W. Landis, and P. Brown. 2001. Developing a regional ecological risk assessment: A case study of a Tasmanian agricultural catchment. *Human and Ecological Risk Assessment* 7(2): 417-439.

Wan M.T., J.-N. Kuo, B. McPherson, and J. Pasternak. 2006. Triazine and metolachlor herbicide residues in farm areas of the Lower Fraser Valley, British Columbia, Canada. *Journal of Environmental Science and Health - Part B Pesticides, Food Contaminants, and Agricultural Wastes* 41(6): 855-867.

Wang J., X. Zhu, X. Zhang, Z. Zhao, H. Liu, R. George, J. Wilson-Rawls, Y. Chang, and Y. Chen. 2011. Disruption of zebrafish (Danio rerio) reproduction upon chronic exposure to TiO2 nanoparticles. *Chemosphere* 83(4): 461-467.

Wang B., G. Yu, J. Huang, H. Hu, Y. Yu, and L. Wang. 2009. Tiered aquatic ecological risk assessment of organochlorine pesticides and their mixture in Jiangsu reach of Huaihe River, China. *Environmental Monitoring and Assessment* 157(1-4): 29-42.

Wang L. 1997. *A Course in Fuzzy Systems and Control*. New Jersey: Prentice Hall.

Weed D.L. 2005. Weight of evidence: A review of concept and methods. *Risk Analysis* 25(6): 1545-1557.

Weyers A., B. Sokull-Klüttgen, T. Knacker, S. Martin, and C.A.M. van Gestel. 2004. Use of Terrestrial Model Ecosystem Data in Environmental Risk Assessment for Industrial Chemicals, Biocides and Plant Protection Products in the EU. *Ecotoxicology* 13(1-2): 163-176.

White R.H., C.H. Stineman, J.M. Symons, P.N. Breysse, S.R. Kim, M.L. Bell, and J.M. Samet. 2008. Premature Mortality in the Kingdom of Saudi Arabia Associated with Particulate Matter Air Pollution from the 1991 Gulf War. *Human and Ecological Risk Assessment* 14: 645-664.

Wiegers J.K., H.M. Feder, L.S. Mortensen, D.G. Shaw, V.J. Wilson, and W.G. Landis. 1998. A regional multiple-stressor rank-based ecological risk assessment for the fjord of Port Valdez, Alaska. *Human and Ecological Risk Assessment* 4(5): 1125-1173.

Williams B.K., R.C. Szaro, and C.D.Shapiro. 2009. *Adaptive Management: The U.S. Department of the Interior Technical Guide*. Washington: Adaptive Management Working Group.

van Wijngaarden R.P.A., J.G.M. Cuppen, G.H.P. Arts, S.J.H. Crum, M.W. Van Den Hoorn, P.J. Van Den Brink, and T.C.M. Brock. 2004. Aquatic risk assessment of a realistic exposure to pesticides used in bulb crops: A microcosm study. *Environmental Toxicology and Chemistry* 23(6): 1479-1498.

Wolfenbarger L.L. and P.R. Phifer. 2000. The ecological risks and benefits of genetically engineered plants. *Science* 290: 2088-2093.

Wolt J.D., P. Kees, A. Raybould, J.W. Fitzpatrick, M. Burachik, A. Gray, S.S. Olin, J. Schiemann, M. Sears, and F. Wu. 2009. Problem formulation in the environmental risk assessment for genetically modified plants. *Transgenic Research* 19(3): 425-436.

Wolt J.D., R.K.D. Peterson, P. Bystrak, T. Meade. 2003. A screening level approach for nontarget insect risk assessment: Transgenic Bt corn pollen and the monarch butterfly (Lepidoptera: Danaidae). *Environmental Entomology* 32(2): 237-246.

World Health Organization. 2008. *Uncertainty and Data Quality in Exposure Assessment*. Geneva: WHO Press.

World Health Organization. 2006. *Health Risk of Particulate Matter from Long-range Trans-boundary Air Pollution*. Bonn, Germany: European Centre for Environment and Health.

Wright-Walters M., C. Volz, E. Talbott, and D. Davis. 2011. An updated weight of evidence approach to the aquatic hazard assessment of Bisphenol A and the derivation a new predicted no effect concentration (Pnec) using a non-parametric methodology. *Science of the Total Environment* 409(4): 676-685.

Wu F. and Y. Tsang. 2004. Second-order Monte Carlo uncertainty/variability analysis using correlated model parameters: application to salmonid embryo survival risk assessment. *Ecological Modelling* 177(3-4): 393-414.

Wynne B. 1992. Uncertainty and environmental learning. Reconceiving science and policy in the preventive paradigm. *Global Environmental Change* 2(2): 111-127.

Yeh S. and M.J. Small. 2002. Incorporating exposure models in probabilistic assessment of the risks of premature mortality from particulate matter. *Journal of Exposure Analysis and Environmental Epidemiology* 12(6): 389-403.

Zadeh L.A. 1996. Fuzzy logic = computing with words. *IEEE Transactions on Fuzzy Systems* 4(2): 103-111.

Zadeh L.A. 1975. The concept of a linguistic variable and its application to approximate reasoning-I. *Information Sciences* 8(3): 199-249.

Zadeh L.A. 1965. Fuzzy sets. *Information and Control* 83: 338-353.

Zalk D., S. Paik, and P. Swuste. 2009. Evaluating the Control Banding Nanotool: a qualitative risk assessment method for controlling nanoparticle exposures. *Journal of Nanoparticle Research* 11(7): 1685-1704.

Zangerl A.R., D. McKenna, C.L. Wraight, M. Carroll, P. Ficarello, R. Warner, and M.R. Berenbaum. 2001. Effects of exposure to event 176 Bacillus thuringiensis corn pollen on monarch and black swallowtail caterpillars under field conditions. *Proceedings of the National Academy of Sciences of the United States of America* 98(21): 11908-11912.

Zhang K.,Y. Pei, and C. Lin. 2010. An investigation of correlations between different environmental assessments and risk assessment. *Procedia Environmental Sciences* 2: 643-64.

Zhang H., G.H. Huang, and G.M. Zeng. 2009. Health risks from arsenic-contaminated soil in Flin Flon-Creighton, Canada: Integrating geostatistical simulation and dose-response model. *Environmental Pollution* 157(8-9): 2413-2420.

Zhao X., Y. Zhang, and S. Li. 2008. Ecological risk assessment of DDT accumulation in aquatic organisms of Taihu Lake, China. *Human and Ecological Risk Assessment* 14(4): 819-834.

Zhu X., L. Zhu, Y. Lang, and Y. Chen. 2008. Oxidative stress and growth inhibition in the freshwater fish Carassius auratus induced by chronic exposure to sublethal fullerene aggregates. *Environmental Toxicology and Chemistry* 27(9): 1979-1985.

Zhu X., L. Zhu, Y. Li, Z. Duan, W. Chen, and P.J.J. 2007. Alvarez. Developmental toxicity in zebrafish (Danio rerio) embryos after exposure to manufactured nanomaterials: Buckminsterfullerene aggregates (nC60) and fullerol. *Environmental Toxicology and Chemistry* 26(5): 976-979.

Zimmer K.E., A.C. Gutleb, S. Ravnum. M. Krayer von Krauss, A.J. Murk, E. Ropstad, J.U. Skaare, G. Sundstøl Eriksen, J.L Lyche, J.G. Koppe, *et al.* 2012. Policy relevant results from an expert elicitation on the health risks of phthalates. *Environmental Health* 11(Suppl. 1): S6.

[This page is intentionally left blank]

Appendix A The 171 sourced articles that comprise the WOE evidence base used in Chapter 4, showing some of the extracted information including the risk domain, the uncertainties identified, the UMTs utilised, and the type of WOE framework (see Section 4.3.1).

ID	Reference	Risk domain	Uncertainty identified	UMT(s) utilised	WOE framework
1	Acosta *et al.* 2010	Ecology	Language: vagueness Language: underspecificity Language: ambiguity	Fuzzy logic Fuzzy logic Fuzzy logic	Semi-quantitative: logic
2	Agüero *et al.* 2008	Toxicology	Data: reliability	Monte-Carlo simulation	Quantitative: computational modelling
3	Ahlers *et al.* 2008	Toxicology	Extrapolation: spatial Extrapolation: interspecies	Multi-criteria decision analysis Multi-criteria decision analysis	Quantitative: multi-criteria decision analysis
4	Alden *et al.* 2005	Toxicology	Data: reliability	No action	Semi-quantitative: logic
5	Alvarez-Guerra, M	Sediment management	Variability: human	Multi-criteria decision analysis	Semi-quantitative: ranking; multi-criteria decision analysis
6	An *et al.* 2007	Microbiology	Variability: natural Data: reliability	Monte-Carlo simulation Monte-Carlo simulation	Quantitative: computational modelling
7	Apitz *et al.* 2007	Sediment management	Data: reliability	No action	Semi-quantitative: logic
8	ApSimon *et al.* 2002	Contamination studies	Data: reliability Model: structure Variability: natural	Monte-Carlo simulation Monte-Carlo simulation Monte-Carlo simulation	Quantitative: computational modelling
9	Arhonditsis *et al.* 2007	Ecology	Model: structure Data: reliability (model input)	Sensitivity analysis; Monte-Carlo simulation Sensitivity analysis; Monte-Carlo simulation	Quantitative: computational modelling
10	Aspinall *et al.* 2003	Volcanology	Decision	Bayesian belief network	Quantitative: statistics
11	Avagliano and Parrella 2009	Contamination studies	Model: output Data: reliability (model input)	Sensitivity analysis Further data collection	Quantitative: computational modelling
12	Avagliano *et al.* 2005	Contamination studies	Data: reliability (model input) Model: output	Sensitivity analysis Sensitivity analysis	Quantitative: computational modelling
13	Babendreier and Castleton 2005	Hazardous materials	Model: structure	Monte-Carlo simulation	Quantitative: computational modelling
14	Baccou *et al.* 2008	Radionuclides	Data: availability (model input)	Monte-Carlo simulation	Quantitative: numerical modelling
15	Barron *et al.* 2004	Toxicology	Data: reliability Extrapolation: interspecies	Latin hypercube sampling Latin hypercube sampling	Quantitative: computational modelling
16	Batley *et al.* 2002	Sediment management	Data: reliability Extrapolation: laboratory Extrapolation: interspecies Extrapolation: intraspecies	Uncertainty factor Uncertainty factor Uncertainty factor Uncertainty factor	Semi-quantitative: logic
17	Batzias and Siontorou 2007	Contamination studies	Data: reliability	Further data collection	Quantitative: computational modelling
18	Baudrit *et al.* 2007	Contamination studies	Variability: natural Data: precision	Monte-Carlo simulation Fuzzy logic	
19	Benekos *et al.* 2007	Contamination studies	Data: reliability (model input) Model: output Data: availability	Monte-Carlo simulation Monte-Carlo simulation Monte-Carlo simulation	Quantitative: computational modelling

20	Benke and Hamilton 2008	Microbiology	Data: reliability (model input)	Monte-Carlo simulation	Quantitative: numerical modelling
21	Bennett *et al.* 2007	Toxicology	Data: reliability	Monte-Carlo simulation	Quantitative: computational modelling
22	Beyer *et al.* 2009	Contamination studies	Data: reliability (model input) Model: structure	Monte-Carlo simulation Model validation	Quantitative: numerical modelling
23	Bittueva *et al.* 2007	Plant Science	Variability: natural	Other (discriminant analysis)	Semi-quantitative: ranking
24	Blazkova and Beven 2004	Hydrology	Data: availability (model input) Model: output	Monte-Carlo simulation Fuzzy logic	Quantitative: computational modelling
25	Borsuk *et al.* 2006	Ecology	Data: reliability (model input) Model: structure Variability: natural	Bayesian belief network Bayesian belief network Latin hypercube sampling	Quantitative: numerical modelling
26	Bosgra *et al.* 2005	Toxicology	Extrapolation: interspecies Extrapolation: intraspecies Extrapolation: quantity Data: reliability (model input)	Monte-Carlo simulation Monte-Carlo simulation Monte-Carlo simulation Monte-Carlo simulation	Quantitative: computational modelling
27	Bosgra *et al.* 2009	Toxicology	Data: reliability (model input) Extrapolation: interspecies Extrapolation: intraspecies	Monte-Carlo simulation Monte-Carlo simulation Monte-Carlo simulation	Quantitative: numerical modelling
28	Brechignnac and Doi 2009	Toxicology	Extrapolation: interspecies Extrapolation: quantity Extrapolation: spatial System: process	Precautionary management Precautionary management Precautionary management Precautionary management	
29	Brouwer and De Blois 2008	Contamination studies	System: process Variability: natural Extrapolation: spatial System: cause	Monte-Carlo simulation Monte-Carlo simulation Monte-Carlo simulation Causal influence	Semi-quantitative: ranking
30	Buekers *et al.* 2009	Toxicology	Data: reliability	Further data collection	Semi-quantitative: logic
31	Burgman *et al.* 1999	Ecology	Variability: natural Model: structure Variability: human	Monte-Carlo simulation Monte-Carlo simulation Other (statistics)	Semi-quantitative: ranking
32	Burton *et al.* 2005	Toxicology	Extrapolation: laboratory Variability: natural	Further data collection Further data collection	Semi-quantitative: logic
33	Caley *et al.* 2006	Biological science	Model: structure System: process	Sensitivity analysis Bootstrapping	Quantitative: statistics
34	Campbell and Longsine 1990	Hazardous materials	Model: output Data: availability (model input) Variability: natural	Monte-Carlo simulation; Latin hypercube sampling Monte-Carlo simulation; Latin hypercube sampling Monte-Carlo simulation; Latin hypercube sampling	Quantitative: computational modelling
35	Cañellas-Boltà *et al.* 2005	Sediment management	System: effect System: process	Precautionary management Precautionary management	Qualitative: listing evidence
36	Carlon *et al.* 2008	Contamination studies	Data: reliability (model input) Extrapolation: spatial	Monte-Carlo simulation Interpolation	Quantitative: computational modelling
37	Carrington *et al.* 1997	Toxicology	Variability: natural	Monte-Carlo simulation	Quantitative: computational modelling

			Model: structure Model: output	Sensitivity analysis Monte-Carlo simulation	
38	Cesar *et al.* 2009	Toxicology	Data: reliability	No action	Quantitative: statistics
39	Chapman 2007	Contamination studies	Variability: natural Data: reliability	Further data collection Further data collection	Semi-quantitative: ranking
40	Chen and Ma 2007	Hazardous materials	Data: reliability (model input) Variability: natural Model: output	Monte-Carlo simulation Monte-Carlo simulation Sensitivity analysis	Quantitative: computational modelling
41	Chen *et al.* 2007	Water quality	Data: reliability Variability: human Data: availability	Bayesian belief network; Expert elicitation Bayesian belief network; Expert elicitation Bayesian belief network; Expert elicitation	Semi-quantitative: logic
42	Chowdhury and Flentje 2002	Geology	Extrapolation: spatial System: process Model: structure	Expert elicitation Expert elicitation Sensitivity analysis	Qualitative: best professional judgement
43	Chowdhury *et al.* 2009	Contamination studies	Variability: natural System: process Data: precision Language: vagueness (linguistic) Language: ambiguity	Latin hypercube sampling Latin hypercube sampling Fuzzy logic Fuzzy logic Fuzzy logic	Quantitative: numerical modelling and
44	Collins *et al.* 2000	Nutrient loading	Data: reliability (model input)	Error propagation; Bootstrapping	Quantitative: statistics
45	Collins *et al.* 2004	Toxicology	Extrapolation: intraspecies Extrapolation: interspecies Extrapolation: temporal	Uncertainty factor Uncertainty factor Uncertainty factor	Qualitative: listing evidence
46	Cothern *et al.* 1986	Contamination studies	Extrapolation: intraspecies Variability: natural Extrapolation: spatial Extrapolation: temporal	Uncertainty factor Uncertainty factor Uncertainty factor Uncertainty factor	Qualitative: listing evidence
47	Crane and MacDonald 2003	Contamination studies	System: cause	Causal influence	Semi-quantitative: logic
48	Critto *et al.* 2007	Contamination studies	Decision	Multi-criteria decision analysis	Quantitative: multi-criteria decision analysis
49	Croke *et al.* 2007	Water management	Variability: human: interpretation	Bayesian belief network	Qualitative: best professional judgement
50	Culp *et al.* 2000	Toxicology	Data: reliability	No action	Qualitative: listing evidence
51	Cupit *et al.* 2002	Toxicology	Decision	No action	Semi-quantitative: ranking
52	Daniels *et al.* 2000	Contamination studies	Data: reliability	Monte-Carlo simulation	Quantitative: numerical modelling
53	de Nazelle and Rodríguez 2009	Contamination studies	Data: reliability (model input)	Monte-Carlo simulation; Sensitivity analysis	Quantitative: computational modelling
54	De Nijs *et al.* 1993	Toxicology	Variability: natural	Uncertainty factor	Quantitative: computational modelling
55	DelValls and Riba 2007	Sediment management	Variability: natural	No action	Semi-quantitative: logic
56	Dey *et al.* 2000	Water quality	Data: availability (model input) Decision	Monte-Carlo simulation Adaptive management	Semi-quantitative: causal criteria
57	Diodato and Ceccarelli 2005	Water management	Extrapolation: spatial	Interpolation	Quantitative: computational modelling and statistics
58	Ducey and Larson 1999	Ecology	Variability: human Decision	Fuzzy logic Multi-criteria decision analysis	Semi-quantitative: ranking
59	Dunham *et al.* 2003	Ecology	Decision	Adaptive management	Semi-quantitative: causal criteria

60	Dussault *et al.* 2008	Toxicology	Data: reliability System: cause	Uncertainty factor Uncertainty factor	Semi-quantitative: causal criteria
61	Echevarria *et al.* 2001	Radiation	Variability: natural	Further data collection	Semi-quantitative: ranking
62	Efroymson *et al.* 2007	Water quality	Extrapolation: laboratory Extrapolation: spatial Extrapolation: interspecies	Error propagation Error propagation Error propagation	Semi-quantitative: causal criteria
63	Enick and Moore 2007	Toxicology	Extrapolation: quantity System: process Data: availability	Uncertainty factor Uncertainty factor Expert elicitation	Qualitative: listing evidence
64	Fewtrell *et al.* 2001	Water quality	Data: reliability	Uncertainty factor	Qualitative: best professional judgement
65	Fiksel 1985	Hazardous materials	Data: availability Model: structure Variability: natural Extrapolation: interspecies Data: precision Extrapolation: temporal Extrapolation: spatial	No action No action No action Sensitivity analysis No action No action No action	Quantitative: computational modelling
66	Filipsson *et al.* 2009	Hazardous materials	Data: reliability Data: reliability (model input)	Confidence interval Other (probability bounds analysis - PBA)	Quantitative: computational modelling and statistics
67	Fischer 2005	Toxicology	System: process	Uncertainty factor	Semi-quantitative: logic
68	Fish *et al.* 2009	Environmental policy	Variability: human: interpretation	Expert elicitation	Qualitative: best professional judgement
69	Forbes and Calow 2002	Ecology	System: cause System: effect	No action No action	Semi-quantitative: causal criteria
70	Fuhrer 2009	Ecology	Model: output Data: reliability (model input)	No action No action	Qualitative: listing evidence
71	Godduhn and Duffy 2003	Toxicology	Extrapolation: interspecies Extrapolation: temporal	Precautionary Management Precautionary Management	Semi-quantitative: causal criteria
72	Golden *et al.* 1997	Toxicology	System: effect	Uncertainty factor	Qualitative: listing evidence
73	Goodman *et al.* 1997	Toxicology	System: process Variability: natural Data: reliability Extrapolation: interspecies	No action No action No action No action	Qualitative: listing evidence
74	Gottschalk *et al.* 2010	Nanotoxicology	Data: reliability (model input) Data: availability System: effect	Sensitivity analysis; Monte-Carlo simulation Sensitivity analysis; Monte-Carlo simulation Sensitivity analysis; Monte-Carlo simulation	Quantitative: computational modelling
75	Greenberg 1997	Toxicology	Extrapolation: intraspecies	Uncertainty factor	Semi-quantitative: logic
76	Griffin *et al.* 1999	Toxicology	Data: reliability (model input)	Monte-Carlo simulation	Quantitative: computational modelling
77	Grist *et al.* 2003	Water quality	Extrapolation: intraspecies Variability: natural Data: reliability	bootstrapping bootstrapping bootstrapping	Quantitative: statistics
78	Gurjar and Mohan 2002	Toxicology	Extrapolation: interspecies Extrapolation: quantity	Uncertainty factor Uncertainty factor	Semi-quantitative: causal criteria
79	Hacon *et al.* 1997	Toxicology	Data: reliability (model input)	Monte-Carlo simulation; Latin hypercube	Quantitative: computational modelling

80	Hamilton *et al.* 2006	Microbiology	Data: availability (model input)	Sensitivity analysis	Quantitative: numerical modelling and statistics
81	Hayes and Landis 2004	Ecology	Extrapolation: spatial Extrapolation: interspecies Data: reliability System: effect System: cause Extrapolation: temporal	Monte-Carlo simulation Monte-Carlo simulation Monte-Carlo simulation Monte-Carlo simulation Monte-Carlo simulation Monte-Carlo simulation	Semi-quantitative: ranking
82	Hays *et al.* 2009	Toxicology	Extrapolation: intraspecies Extrapolation: interspecies	Uncertainty factor Uncertainty factor	Semi-quantitative: logic
83	Henning-de Jong *et al.* 2008	Toxicology	Data: reliability (model input)	Monte-Carlo simulation; Latin hypercube sampling	Semi-quantitative: ranking and computational modelling
84	Hughes *et al.* 2003	Toxicology	Extrapolation: interspecies System: cause Extrapolation: temporal Extrapolation: intraspecies System: effect	No action No action No action No action No action	Qualitative: listing evidence
85	Hung *et al.* 2009	Hazardous materials	Data: availability (model input)	Monte-Carlo simulation	Quantitative: computational modelling
86	Huysmans *et al.* 2006	Contamination studies	Data: availability Model: output	Monte-Carlo simulation; Sensitivity analysis Monte-Carlo simulation; Sensitivity analysis	Quantitative: computational modelling
87	Jackson *et al.* 2004	Toxicology	Extrapolation: spatial	No action	Quantitative: computational modelling
88	Jones *et al.* 2009	Toxicology	Extrapolation: interspecies	Monte-Carlo simulation	Quantitative: computational modelling
89	Kaloudis *et al.* 2005	Wildfire	Extrapolation: temporal Extrapolation: spatial Data: reliability Data: availability (model input) Model: output	Fuzzy logic Fuzzy logic Fuzzy logic Fuzzy logic Fuzzy logic	Quantitative: computational modelling
90	Kandlikar *et al.* 2007	Nanotoxicology	Variability: human Extrapolation: intraspecies System: effect Extrapolation: interspecies Variability: natural System: process	Expert elicitation; Probability density function; Bayesian belief network Expert elicitation; Probability density function; Bayesian belief network Expert elicitation; Probability density function; Bayesian belief network Expert elicitation; Probability density function; Bayesian belief network Expert elicitation; Probability density function; Bayesian belief network Expert elicitation; Probability density function; Bayesian belief network	Qualitative: best professional judgement
91	Kapo and Burton 2006	Toxicology	Extrapolation: spatial	Sensitivity analysis	Quantitative: computational modelling
92	Keiter *et al.* 2009	Toxicology	Data: reliability (empirical) Variability: natural Language: ambiguity Language: vagueness (linguistic) Language: underspecificity	Fuzzy logic Fuzzy logic Fuzzy logic Fuzzy logic Fuzzy logic	Semi-quantitative: logic

93	Kelly *et al.* 2009	Toxicology	Extrapolation: intraspecies Extrapolation: spatial Extrapolation: temporal	Further data collection Further data collection Further data collection	Semi-quantitative: logic
94	Kentel and Aral 2007	Toxicology	Variability: natural Data: availability Data: reliability System: process	Monte-Carlo simulation Fuzzy logic Fuzzy logic Fuzzy logic	Quantitative: computational modelling
95	King and Richardson 2003	Water quality	System: process	Bootstrapping	Quantitative: statistics
96	Klier *et al.* 2008	Plant Science	Data: reliability (model input)	Latin hypercube sampling	Quantitative: computational modelling
97	Kooistra *et al.* 2005	Ecology	Data: availability (model input) Extrapolation: spatial	Monte-Carlo simulation Interpolation	Quantitative: computational modelling
98	Krayer von Krauss *et al.* 2004	Plant Science	System: process Data: availability Model: output	Expert elicitation Expert elicitation Expert elicitation	Qualitative: best professional judgement
99	Kumar *et al.* 2009	Toxicology	Variability: natural Data: reliability Data: precision Extrapolation: spatial Extrapolation: temporal	Fuzzy-stochastic Fuzzy-stochastic Fuzzy-stochastic Fuzzy-stochastic Fuzzy-stochastic	Quantitative: computational modelling
100	Landis *et al.* 2004	Ecology	Variability: natural Extrapolation: temporal	Further data collection Further data collection	Semi-quantitative: ranking
101	Lee *et al.* 2008	Toxicology	Data: availability (model input) Extrapolation: spatial	Monte-Carlo simulation Interpolation	Semi-quantitative: causal criteria
102	Lemke and Bahrou 2009	Contamination studies	Extrapolation: interspecies Data: reliability	Monte-Carlo simulation Monte-Carlo simulation	Quantitative: computational modelling
103	Li *et al.* 2006	Contamination studies	Data: availability Extrapolation: interspecies	Fuzzy logic Fuzzy logic	Quantitative: computational modelling
104	Li *et al.* 2007	Contamination studies	Language: vagueness Language: ambiguity Data: availability	Fuzzy-stochastic Fuzzy-stochastic Fuzzy-stochastic	Quant computational modelling
105	Li *et al.* 2008	Contamination studies	Language: vagueness (linguistic) Language: underspecificity Data: availability	Fuzzy-stochastic Fuzzy-stochastic Fuzzy-stochastic	Quantitative: computational modelling
106	Liao and Chou 2005	Toxicology	System: process Data: availability	Monte-Carlo simulation; Sensitivity analysis; Model validation Monte-Carlo simulation; Sensitivity analysis; Model validation	Quantitative: computational modelling
107	Lindenschmidt *et al.* 2008	Water management	Data: reliability (model input)	Sensitivity analysis; Monte-Carlo simulation	Quantitative: computational modelling
108	Linkov and Burmistrov 2005	Ecology	Data: reliability (model input)	Monte-Carlo simulation	Quantitative: computational modelling
109	Linkov *et al.* 2001	Sediment management	Data: reliability (model input)	Latin hypercube sampling	Quantitative: computational modelling
110	Liu *et al.* 2007	Contamination studies	Data: reliability (model input)	Monte-Carlo simulation	Quantitative: computational modelling
111	Loos *et al.* 2009	Toxicology	System: effect	Further data collection	Quantitative: computational modelling
112	Lu *et al.* 2003	Ecology	Extrapolation: interspecies Data: reliability (model input)	Uncertainty factor Uncertainty factor	Quantitative: computational modelling

113	Ma and van der Voet 1993	Toxicology	Data: availability (model input)	Error propagation	Quantitative: numerical modelling
114	Ma 2002	Contamination studies	Data: reliability (model input)	Monte-Carlo simulation; Other (statistics)	Quantitative: computational modelling
115	Matson *et al.* 2009	Toxicology	Extrapolation: laboratory	Further data collection	Semi-quantitative: ranking
116	Maxim and McConnell 2001	Toxicology	Extrapolation: interspecies	Uncertainty factor	Semi-quantitative: ranking
117	Maxwell and Kastenberg 1999	Contamination studies	Data: reliability System: process Extrapolation: intraspecies	Monte-Carlo simulation Monte-Carlo simulation Monte-Carlo simulation	Quantitative: computational modelling
118	Maycock and Benford 2007	Toxicology	Extrapolation: interspecies Extrapolation: intraspecies Data: availability	Uncertainty factor Uncertainty factor Uncertainty factor	Semi-quantitative: causal criteria
119	McDonald and Wilcockson 2003	Toxicology	Extrapolation: laboratory Extrapolation: interspecies	Uncertainty factor Uncertainty factor	Semi-quantitative: logic
120	Meek and Hughes 1995	Toxicology	Extrapolation: interspecies Extrapolation: quantity	Uncertainty factor Uncertainty factor	Semi-quantitative: ranking
121	Meek *et al.* 2002	Toxicology	Extrapolation: intraspecies Extrapolation: interspecies Extrapolation: spatial Data: reliability Extrapolation: quantity	Uncertainty factor Uncertainty factor Uncertainty factor Uncertainty factor Uncertainty factor	Semi-quantitative: logic
122	Meyer *et al.* 2009	Water management	Data: reliability	Multi-criteria decision analysis	Quantitative: multi-criteria decision
123	Mugglestone *et al.* 2001	Water quality	Model: output	Confidence interval; Error propagation	Qualitative: listing evidence
124	Mukhtasor *et al.* 2004	Water quality	Data: availability Data: reliability (model input)	Monte-Carlo simulation; Latin hypercube sampling Monte-Carlo simulation; Latin hypercube sampling	Quantitative: computational modelling
125	Naito *et al.* 2006	Toxicology	System: process Extrapolation: laboratory System: effect	Uncertainty factor Uncertainty factor Uncertainty factor	Semi-quantitative: logic
126	Nasiri *et al.* 2002	Contamination studies	Variability: human Language: ambiguity Language: vagueness (linguistic) Language: underspecificity	Fuzzy logic Fuzzy logic Fuzzy logic Fuzzy logic	Semi-quantitative: ranking
127	Nazir *et al.* 2008	Contamination studies	Data: reliability (model input)	Sensitivity analysis; Bootstrapping	Quantitative: numerical modelling
128	Neuhäuser and Terhorst 2007	Geology	Data: availability (model input)	Uncertainty factor	Semi-quantitative: causal criteria
129	Oughton *et al.* 2008	Radiation	Data: reliability Data: precision Variability: natural Model: output Model: structure Extrapolation: temporal Extrapolation: spatial	Probability density function; Expert elicitation; No action; Uncertainty factor; Probability density function; Expert elicitation; No action; Uncertainty factor; Probability density function; Expert elicitation; No action; Uncertainty factor; Sensitivity analysis Sensitivity analysis Sensitivity analysis Sensitivity analysis	Quantitative: numerical modelling

			System: process	Other (scenario analysis)	
130	Park *et al.* 2008	Toxicology	Data: reliability (model input)	Latin hypercube sampling	Quantitative: computational modelling
131	Pascoe *et al.* 1993	Contamination studies	Extrapolation: spatial	Further data collection	Semi-quantitative: causal criteria
			Extrapolation: temporal	Further data collection	
			Extrapolation: intraspecies	Further data collection	
			Extrapolation: interspecies	Further data collection	
			System: cause	Further data collection	
132	Persson and Destouni 2009	Contamination studies	Data: reliability (model input)	Probability density function	Quantitative: computational modelling
			Variability: natural	Probability density function	
			Data: availability	Probability density function	
			Extrapolation: spatial	Probability density function	
			Model: structure	Probability density function	
133	Phillips *et al.* 2008	Toxicology	Data: availability	Uncertainty factor	Semi-quantitative: logic
			System: effect	Uncertainty factor	
			System: cause	Uncertainty factor	
			Extrapolation: interspecies	Uncertainty factor	
			Extrapolation: intraspecies	Uncertainty factor	
			Extrapolation: quantity	Uncertainty factor	
			Extrapolation: temporal	Uncertainty factor	
134	Pollino *et al.* 2007	Hydrology	System: process	Bayesian belief network	Semi-quantitative: logic
			Data: reliability (model input)	Other (entropy)	
135	Proctor *et al.* 2002	Contamination studies	Data: reliability	Probability density function; Monte-Carlo simulation	Quantitative: computational modelling
136	Prudhomme *et al.* 2003	Climate change	Data: reliability (model input)	Further data collection	Quantitative: computational modelling
			Variability: natural	Further data collection	
			Model: structure	No action	
			Extrapolation: temporal	Monte-Carlo simulation; No action	
			Extrapolation: spatial	Monte-Carlo simulation; No action	
137	Qin and Huang 2009	Contamination studies	Extrapolation: spatial	Fuzzy-stochastic	Quantitative: computational modelling
			Data: availability	Fuzzy-stochastic	
			Data: reliability (model input)	Fuzzy-stochastic	
138	Ranke 2002	Toxicology	Data: reliability (model input)	Monte-Carlo simulation	Quantitative: numerical modelling
			Model: structure	Monte-Carlo simulation	
139	Rutherford *et al.* 2003	Toxicology	Extrapolation: laboratory	Uncertainty factor	Semi-quantitative: ranking
			Extrapolation: interspecies	Uncertainty factor	
			Extrapolation: intraspecies	Uncertainty factor	
			System: cause	Uncertainty factor	
			Data: availability	Uncertainty factor	
140	Sander and Öberg 2006	Contamination studies	Data: precision (model input)	Probability density function; Monte-Carlo simulation	Quantitative: numerical modelling
			Model: output	Probability density function; Monte-Carlo simulation	
141	Sanderson *et al.* 2006	Toxicology	Extrapolation: laboratory	Further data collection	Semi-quantitative: ranking
			Extrapolation: quantity	Further data collection	

142	Sanderson *et al.* 2007	Toxicology	Extrapolation: laboratory	Further data collection	Semi-quantitative: ranking
143	Scherm 2000	Plant science	Data: availability (model input)	Fuzzy logic	Semi-quantitative: numerical modelling
			Variability: natural	Fuzzy logic	
144	Schoeny *et al.* 2006	Water quality	System: process	Uncertainty factor	Qualitative: listing evidence
145	Schwartz *et al.* 2000	Toxicology	Data: reliability (model input)	Monte-Carlo simulation	Quantitative: numerical modelling
			Variability: natural	Monte-Carlo simulation	
146	Scott *et al.* 2005	Hazardous materials	Data: reliability (model input)	Monte-Carlo simulation	Quantitative: computational modelling
			Data: availability	No action	
			Variability: natural	No action	
147	Shakhawat *et al.* 2006	Water quality	Data: reliability (model input)	Fuzzy logic	Quantitative: numerical modelling
			Data: precision	Fuzzy logic	
			Model: structure	Fuzzy logic	
148	Smith *et al.* 2007	Toxicology	Data: availability	Other (statistics)	Qualitative: listing evidence
			Data: reliability	Hazard quotient	
149	Smith *et al.* 2009	Toxicology	Data: availability	Monte-Carlo simulation; Sensitivity analysis	Quantitative: numerical modelling
			Data: reliability (model input)	Monte-Carlo simulation; Sensitivity analysis	
150	Staples *et al.* 2002	Toxicology	Extrapolation: laboratory	Uncertainty factor	Semi-quantitative: logic
151	Stevens *et al.* 2007	Contamination studies	Data: availability (model input)	Probability density function	Quantitative: computational modelling
152	Teunis *et al.* 1997	Water quality	Data: reliability	Monte-Carlo simulation	Quantitative: numerical modelling
153	Therriault and Herborg 2008	Ecology	System: effect	Expert elicitation	Qualitative: best professional judgement
154	Thorsen *et al.* 2006	Contamination studies	Data: reliability (model input)	Monte-Carlo simulation	Quantitative: computational modelling
155	Tillman and Weaver 2006	Contamination studies	Data: reliability (model input)	Sensitivity analysis	Quantitative: computational modelling
			Model: output	Sensitivity analysis	
156	Tsuji *et al.* 2004	Toxicology	Data: reliability	Uncertainty factor	Qualitative: listing evidence
			Extrapolation: quantity	Uncertainty factor	
157	Twining and Cameron 1997	Toxicology	Data: availability	Confidence interval	Semi-quantitative: logic
158	Vallack *et al.* 1998	Toxicology	System: effect	Expert elicitation	Qualitative: best professional judgement
159	van den Brink *et al.* 2008	Contamination studies	Data: reliability	Monte-Carlo simulation; Latin hypercube sampling	Quantitative: computational modelling
			Model: output	Sensitivity analysis	
160	van Sprang *et al.* 2009	Toxicology	Extrapolation: interspecies	Further data collection	Quantitative: computational modelling
			Variability: natural	Bootstrapping; Monte-Carlo simulation	
161	Verdonck *et al.* 2008	Environmental policy	Data: reliability (model input)	Further data collection; Uncertainty factor;	Quantitative: computational modelling
			Data: availability	Further data collection; Uncertainty factor; Monte-Carlo simulation	
			Extrapolation: interspecies	Further data collection; Uncertainty factor; Monte-Carlo simulation	
			Variability: natural	Further data collection; Uncertainty factor; Monte-Carlo simulation	
			Extrapolation: intraspecies	Further data collection; Uncertainty factor; Monte-Carlo simulation	
162	Vu *et al.* 2006	Contamination studies	Data: reliability (model input)	Monte-Carlo simulation	Quantitative: numerical modelling
163	Walker *et al.* 2001	Ecology	System: cause	Sensitivity analysis	Semi-quantitative: ranking
164	Wang *et al.* 2009	Water quality	Data: reliability	Hazard quotient; Monte-Carlo simulation	Quantitative: numerical modelling
			System: effect	Hazard quotient; Monte-Carlo simulation	

			Data: availability	Hazard quotient; Monte-Carlo simulation	
			Extrapolation: laboratory	Hazard quotient; Monte-Carlo simulation	
			Variability: natural	Hazard quotient; Monte-Carlo simulation	
			Extrapolation: intraspecies	Hazard quotient; Monte-Carlo simulation	
165	Weyers *et al.* 2004	Toxicology	Extrapolation: laboratory	Uncertainty factor	Semi-quantitative: logic
			Variability: natural	Uncertainty factor	
			System: process	Uncertainty factor	
			System: cause	Uncertainty factor	
			System: effect	Uncertainty factor	
166	Wiegers *et al.* 1998	Ecology	Data: availability	Sensitivity analysis	Semi-quantitative: ranking
			System: process	Sensitivity analysis	
			Language: ambiguity	Sensitivity analysis	
			Variability: natural	Sensitivity analysis	
			Variability: human	Sensitivity analysis	
167	Wright-Walters *et al.* 2011	Ecology	Extrapolation: interspecies	No action	Semi-quantitative: logic
			Data: reliability	No action	
168	Wu and Tsang 2004	Ecology	System: process	Monte-Carlo simulation	Quantitative: numerical modelling
			Variability: natural	Monte-Carlo simulation	
			Data: reliability (model input)	Monte-Carlo simulation	
169	Xiao *et al.* 2008	Contamination studies	Data: availability (model input)	Monte-Carlo simulation	Quantitative: numerical modelling
170	Zalk *et al.* 2009	Nanotoxicology	System: cause	Uncertainty factor	Semi-quantitative: ranking
			System: effect	Uncertainty factor	
			System: process	Uncertainty factor	
171	Zhang *et al.* 2009	Toxicology	Extrapolation: spatial	Interpolation	Quantitative: computational modelling
			Extrapolation: temporal	Further data collection	
			Variability: natural	Further data collection	

Appendix B Content of the 'introduction' section of the elicitation system distributed to experts participating in the case study of potential agricultural chemical pesticide risk to surface water organisms (see Section 5.2.4).

Organisation of this workbook

Pre-task information:

Tab 1 – Introductory information; Tab 2 – Instructions and an example;

Tasks to be completed - Assessing the uncertainty in:

Tab 3 – Problem formulation; Tab 4 – Exposure assessment; Tab 5 – Effects assessment; Tab 6 – Risk characterisation

An overview

*** The purpose of the research** is to elicit views on uncertainty in Environmental Risk Assessments (ERAs) as a means to help inform and improve the uncertainty identification process.

*** The purpose of this exercise** is to elicit your views on uncertainty relating to critical aspects of ERAs. The system that will be used to do this is based on a (validated) generic ERA template, and is populated with information relating to *potential agricultural chemical pesticide risk to surface water organisms*.

*** What you need to do:** I would like you to consider the uncertainty (see uncertainty information below) that is associated with each ERA task/process to be evaluated in the ERA template. The domain-specific information is there to help contextualise the generic aspects.

*** How long it will take:** If you have some uncertainty-based experience it should take no more than 20-30 minutes. Otherwise, it should take no more than 30-40 minutes. You don't have to complete it all in one go.

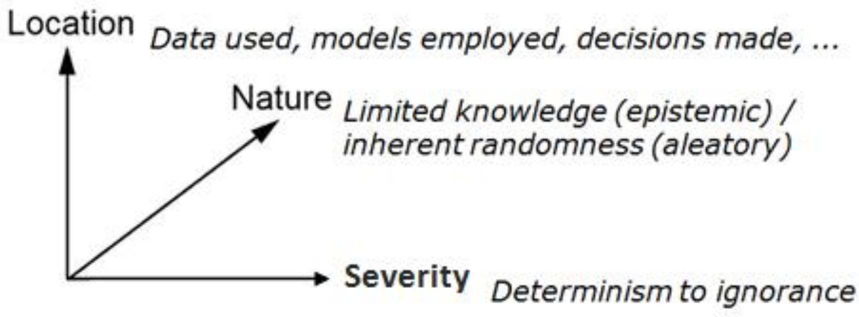
Uncertainty - Some background information

Environmental uncertainty is characterised by three different dimensions (see figure below). These dimensions specifically relate to:

* the ***severity*** of the uncertainty (how bad it is) ranging from deterministic treatment at one end of the spectrum to indeterminacy (i.e. ignorance) at the other;

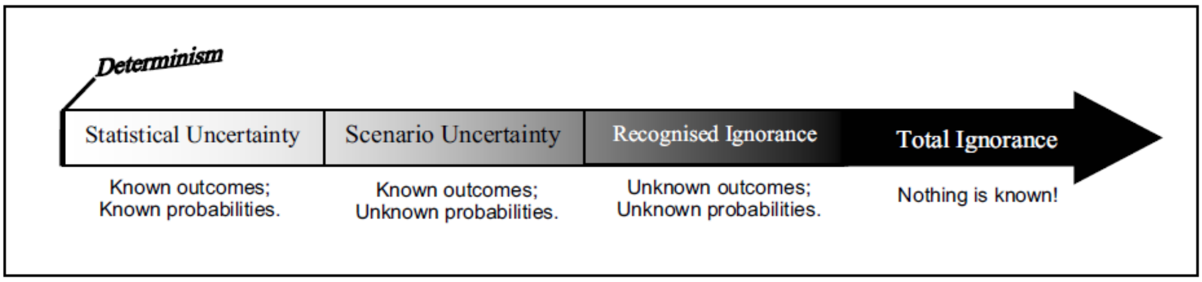
* the ***nature*** of the uncertainty, either aleatory (the randomness of natural systems and their components), epistemic (limitations in our own knowledge), or a combination of both;

* the ***location*** of the uncertainty, which describes where the uncertainty manifests in applied situations;



Dimension 1: The *severity* of uncertainty

The severity of the uncertainty that is associated with an aspect of a system of interest describes how bad that uncertainty is. It is described on a scale from deterministic understanding of the uncertain element through to ignorance to the uncertain element (Figure below).



Uncertainty severity levels when defined through knowledge about likelihoods and knowledge about outcomes (after Walker *et al.* 2003)

Dimension 2: The *nature* of uncertainty

Aleatory uncertainty represents the inherent randomness displayed in human and natural systems. As increasing the knowledge-base associated with these interactions will do nothing to negate their existence, aleatory uncertainty cannot be reduced. However, additional research may help to better understand the complexities of the system(s) of interest.

Epistemic uncertainty represents the imperfection of knowledge concerning a system of interest. In contrast to aleatory uncertainty, epistemic uncertainties can be quantified, reduced, and possibly eliminated, although this depends on the specific situation.

A combination of **both** forms of uncertainty is also a possibility, where it is hard to isolate the individual (aleatory and epistemic) uncertainties.

<u>Dimension 3: The *location* of uncertainty</u>

The location describes the place in the system of interest in which the uncertainty exists. The typology in the table below provides top-level and sub-level examples of the location-based uncertainty that you should be considering throughout this exercise.

Nature	Location	Sub-location	Definition
Epistemic	Data	Availability	referring to the incompleteness, scarcity, or absence of data
		Precision	concerning the lack of accuracy or precision in obtained data
		Reliability	reflecting its trustworthiness i.e. data is erroneous for some specified reason
	Language	Ambiguity	where multiple meanings are possible
		Underspecificity	where meanings are not exact
		Vagueness	where meanings are not clear and understandable
	System	Cause	concerning a lack of clarity regarding the source(s) of harm
		Effect	relating to the influence a particular stressor (source) has upon the receptor(s)
		Process	where the risks are not understood or a process vital to a successful assessment is not identified
Aleatory	Variability	Human	which exists through intentionally biased and subjective human actions
		Natural	which pertains to the stochastic traits of natural systems
	Extrapolation	Intraspecies	where information specific to members of a species is used to represent other members of the same species
		Interspecies	where information specific to members of a species is used to represent members of a different species
		Laboratory	where information specific to laboratory conditions is used to represent real-world scenarios
		Quantity	where information specific to one quantity is used to represent another
		Spatial	where information specific to one spatial scale is used to represent another
		Temporal	where information specific to one timescale is used to represent another
Both	Model	Structure	concerning the representation of real-world processes in model form
		Output	reflecting the level of confidence in the produced results
	Decision	Decision	where doubt surrounds an optimal course of action, often in the face of differing objectives. Decision uncertainty is potentially comprised of all other identified uncertainties

<u>The ERA process</u>

A generic ERA template has been created, the components of which are used as the basis for this elicitation exercise. The template has been created using existing grey and peer-reviewed documents – it does not represent anything new. It is focussed on handling **potential risks to environmental quality, ecological assets, and human-health that originate through human actions (e.g. GM crops, chemicals, nanomaterials)** rather than those that occur naturally and are exacerbated by human actions (e.g. flooding, climate change). The template has been through two rounds of validation (involving 50+ experts) and has been updated accordingly.

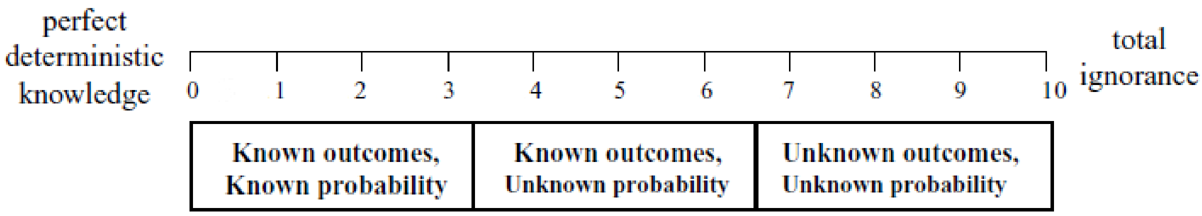
Appendix C Content of the 'instructions' section of the elicitation system distributed to experts participating in the case study of potential agricultural chemical pesticide risk to surface water organisms (see Section 5.2.4).

<u>Instructions on how to complete this elicitation exercise</u>

The four workbook tabs following this one contain the sections for you to complete. Each tab represents a different stage of the ERA process and contains tasks that are specific to that stage. Relevant information from the domain-specific example of *potential agricultural chemical pesticide risk to surface water organisms* is used to contextualise those generic tasks. For each of the stages of the ERA template, I would like for you to **give your views on the three dimensions of uncertainty that are associated with performing the listed tasks.**

*<u>How to assess the **severity** of uncertainty</u>*

In order to judge the severity of uncertainty associated with performing each task, please use the [0 - 10] scale shown below, where 0 represents a perfect deterministic understanding of the task (i.e. there is **no uncertainty**) and 10 represents total ignorance of the task (i.e. there is **total uncertainty**). The higher the number, the more severe the uncertainty.



(after Krayer von Kraus *et al.* 2004)

In order to judge the nature(s) and location(s) of uncertainty associated with carrying out each task, please use the typology below as a guide.

Nature	Location	Sub-location	Definition
Epistemic	Data	Availability	referring to the incompleteness, scarcity, or absence of data
		Precision	concerning the lack of accuracy or precision in obtained data
		Reliability	reflecting its trustworthiness i.e. data is erroneous for some specified reason
	Language	Ambiguity	where multiple meanings are possible
		Underspecificity	where meanings are not exact
		Vagueness	where meanings are not clear and understandable
	System	Cause	concerning a lack of clarity regarding the source(s) of harm
		Effect	relating to the influence a particular stressor (source) has upon the receptor(s)
		Process	where the risks are not understood or a process vital to a successful assessment is not identified
Aleatory	Variability	Human	which exists through intentionally biased and subjective human actions
		Natural	which pertains to the stochastic traits of natural systems
	Extrapolation	Intraspecies	where information specific to members of a species is used to represent other members of the same species
		Interspecies	where information specific to members of a species is used to represent members of a different species
		Laboratory	where information specific to laboratory conditions is used to represent real-world scenarios
		Quantity	where information specific to one quantity is used to represent another
		Spatial	where information specific to one spatial scale is used to represent another
		Temporal	where information specific to one timescale is used to represent another
Both	Model	Structure	concerning the representation of real-world processes in model form
		Output	reflecting the level of confidence in the produced results
	Decision	Decision	where doubt surrounds an optimal course of action, often in the face of differing objectives. Decision uncertainty is potentially comprised of all other identified uncertainties

A worked example

Please have a go at the following pre-elicitation practice exercise. It will help to affirm the uncertainty-based concepts that you will need throughout the elicitation and get you used to the format of the workbook tabs.

Example topic: The introduction of a DNA vaccine in aquaculture (Gillund *et al.* 2008).

Overview: Aquaculture is becoming an increasingly more important source of fish and shellfish products. Advancements in vaccination have played an important role in the expansion of cultivation of high quality fish species like salmonids. It has however been difficult to develop traditional vaccines to protect against certain viral and parasitic diseases, and DNA vaccines are considered a promising solution to this problem.

Potential concerns: Based on a review of the benefits and risks of DNA vaccines Gillund *et al.* (2008) developed a figure (see below) illustrating relevant aspects to consider when explaining potential beneficial and adverse consequences of introducing DNA vaccines in aquaculture. The figure features three pathways of the fate of the DNA vaccine: 1) immune response, 2) distribution of the DNA after injection and 3) potential environmental release:

1) ***Immune response*** is the response within the receptor that is induced by the vaccine. We distinguish between the **intended** immune response, which is the immune response the DNA vaccine is developed to enforce, and **unintended** immune responses.

2) ***Distribution of DNA***: The DNA vaccine will distribute within the host. Persisting **Intact DNA** taken up by host cells may express its DNA encoded gene, or **Degraded DNA** may integrate into the chromosomal DNA of the fish, affecting future generations.

3) ***Environmental release***: DNA from the vaccines could unintentionally be distributed over vast areas.

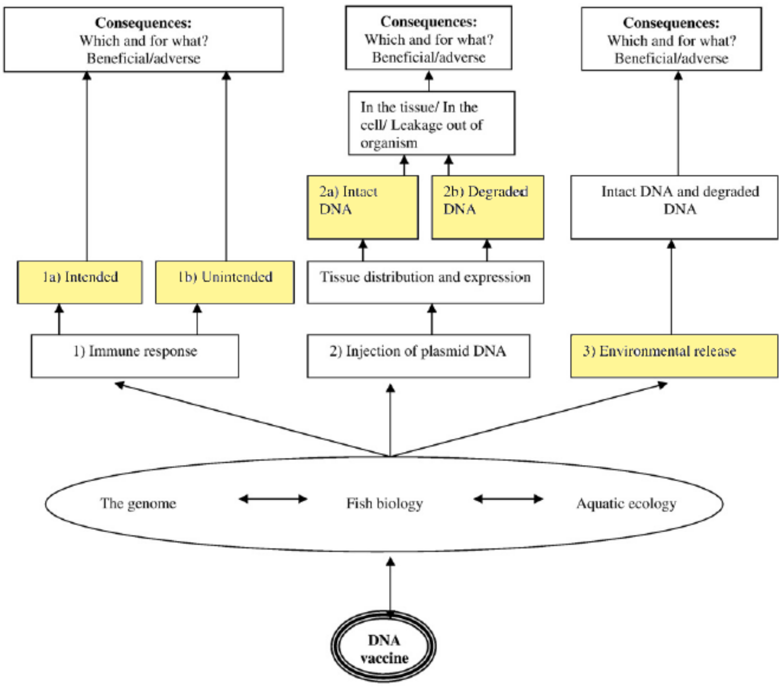


Fig. 2 – Potential beneficial and adverse consequences of introducing DNA vaccines in aquaculture.

<u>The exercise</u>: Consider the *severity*, *nature*, and potential *locations* of uncertainty associated with determining the following key aspects of DNA vaccination of fish (highlighted yellow in the figure opposite):

1) Intended immune responses

2) Unintended immune responses

3) Distribution of intact DNA in tissue (after injection)

4) Distribution of degraded DNA in tissue (after injection)

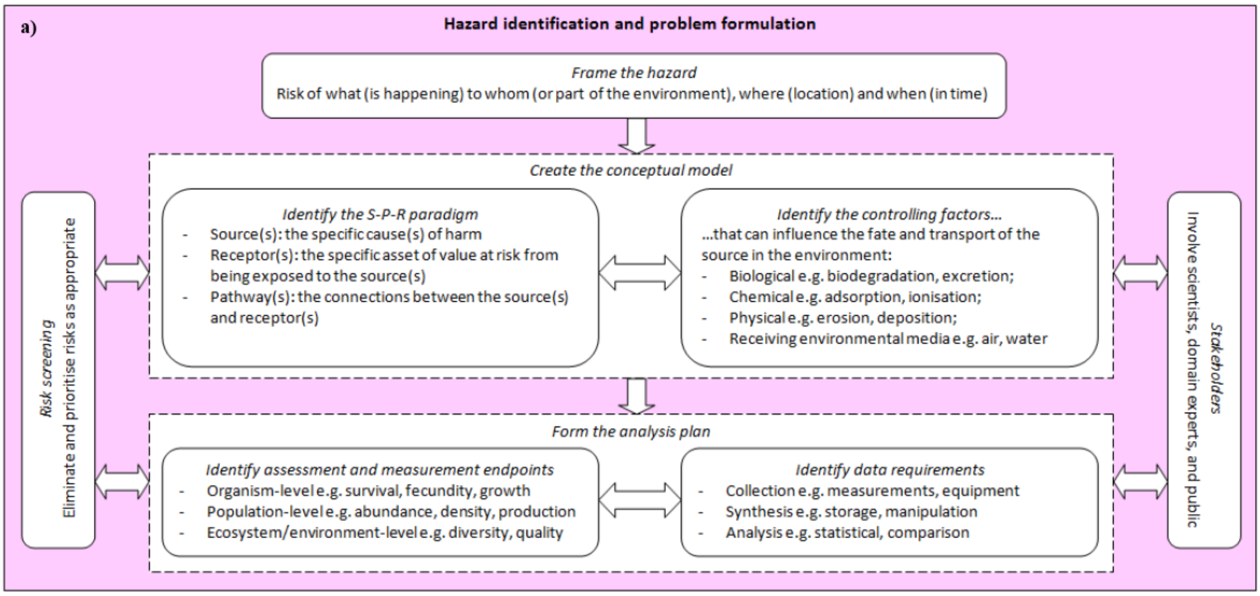
5) Potential for environmental release

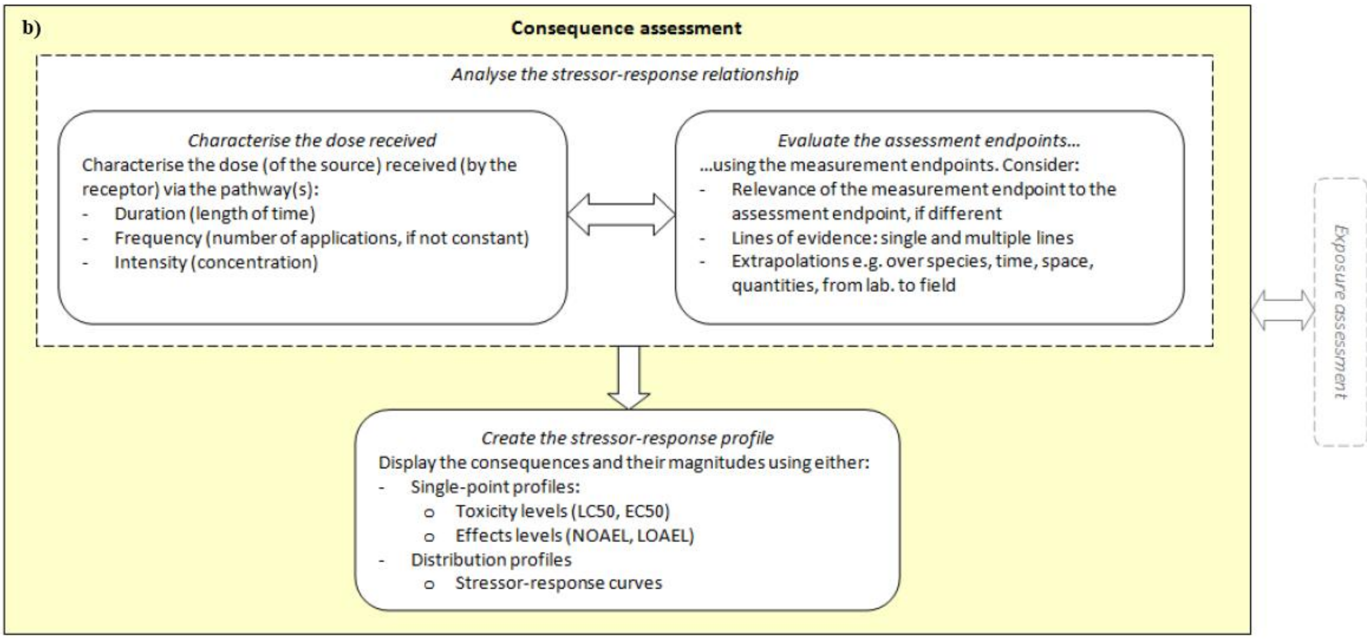
Please begin the short practice exercise below.

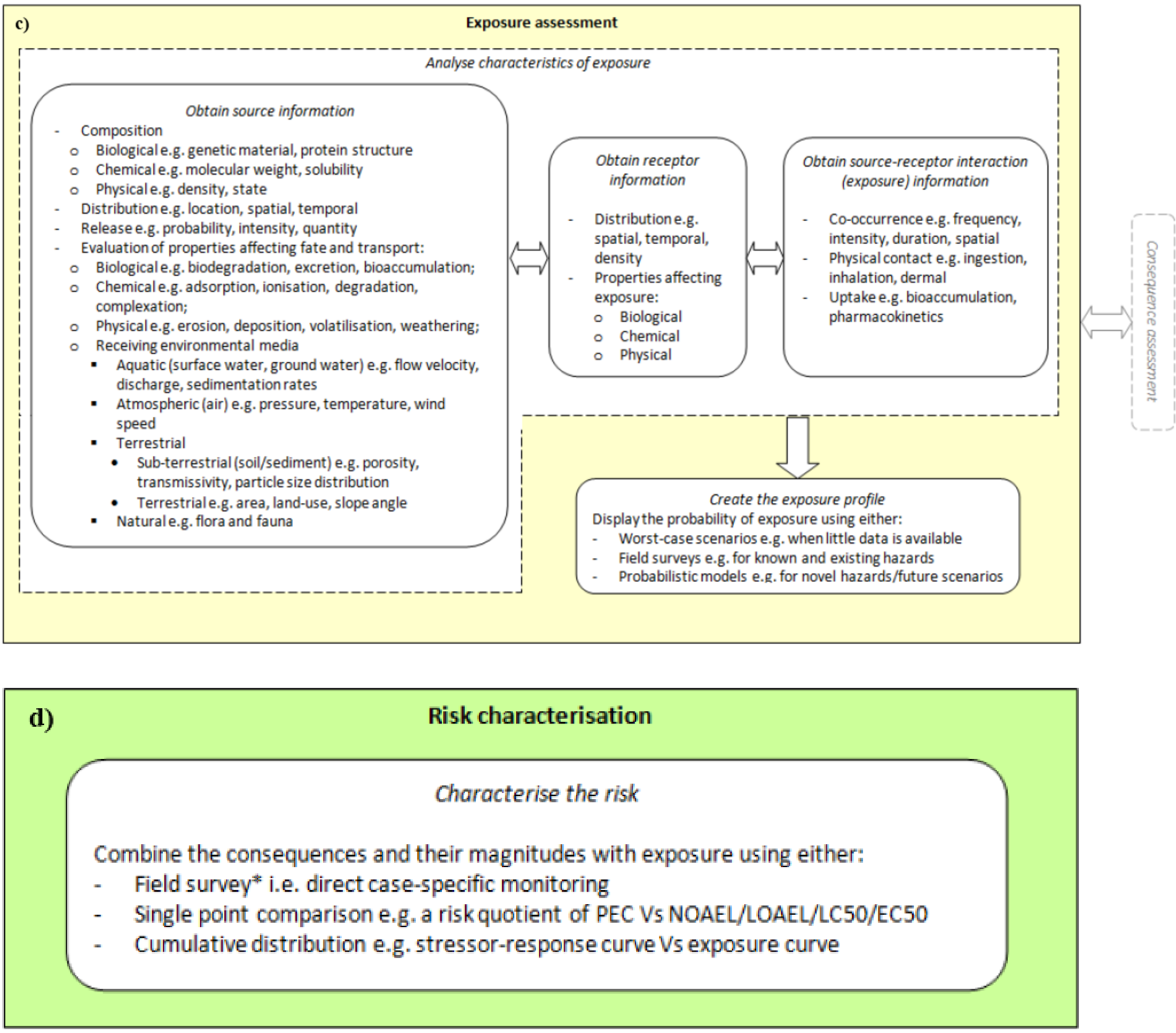
TIP: Click on the red triangles for further guidance.

ERA task to be completed	Domain-specific ERA value(s)	Is the task necessary?	Severity of uncertainty associated with task	Nature of uncertainty associated with task	Location of uncertainty associated with task	Additional comments about the uncertainty, task, etc
1. Determining the immune responses within the receptor						
Determining the intended immune responses of the vaccine	Positive effects of the vaccine in the fish e.g. protection against disease.	☐ Task is not valid	0	☐ Epsitemic ☐ Aleatory ☐ Both	☐ Data ☐ Language ☐ System understanding ☐ Variability ☐ Extrapolation ☐ Model ☐ Decision	
Determining the unintended immune responses of the vaccine	Negative effects of the vaccine in the fish e.g. decreased fish welfare or loss in the aquaculture production.	☐ Task is not valid	0	☐ Epsitemic ☐ Aleatory ☐ Both	☐ Data ☐ Language ☐ System understanding ☐ Variability ☐ Extrapolation ☐ Model ☐ Decision	
2. Determining the distribution of DNA within the tissue of the receptor						
Determining the distribution of intact DNA in the receptor	Using techniques to determine levels of intact DNA in fish injected with vaccine e.g. altered gene expression.	☐ Task is not valid	0	☐ Epsitemic ☐ Aleatory ☐ Both	☐ Data ☐ Language ☐ System understanding ☐ Variability ☐ Extrapolation ☐ Model ☐ Decision	
Determining the distribution of degraded DNA in the receptor	Using techniques to determine levels of degraded DNA in fish injected with vaccine e.g. mutations in offspring.	☐ Task is not valid	0	☐ Epsitemic ☐ Aleatory ☐ Both	☐ Data ☐ Language ☐ System understanding ☐ Variability ☐ Extrapolation ☐ Model ☐ Decision	
3. Determining the potential for environmental release	From e.g. leakage/spillage of intact material or phyla transfer of degraded material.	☐ Task is not valid	0	☐ Epsitemic ☐ Aleatory ☐ Both	☐ Data ☐ Language ☐ System understanding ☐ Variability ☐ Extrapolation ☐ Model ☐ Decision	

Appendix D The generic ERA template, version 1, created through the interrogation of several published peer-reviewed and grey literature sources, describing the aspects within the stages of: **a)** hazard identification and problem formulation; **b)** consequence assessment; **c)** exposure assessment; and **d)** risk characterisation (see Section 5.3.2).

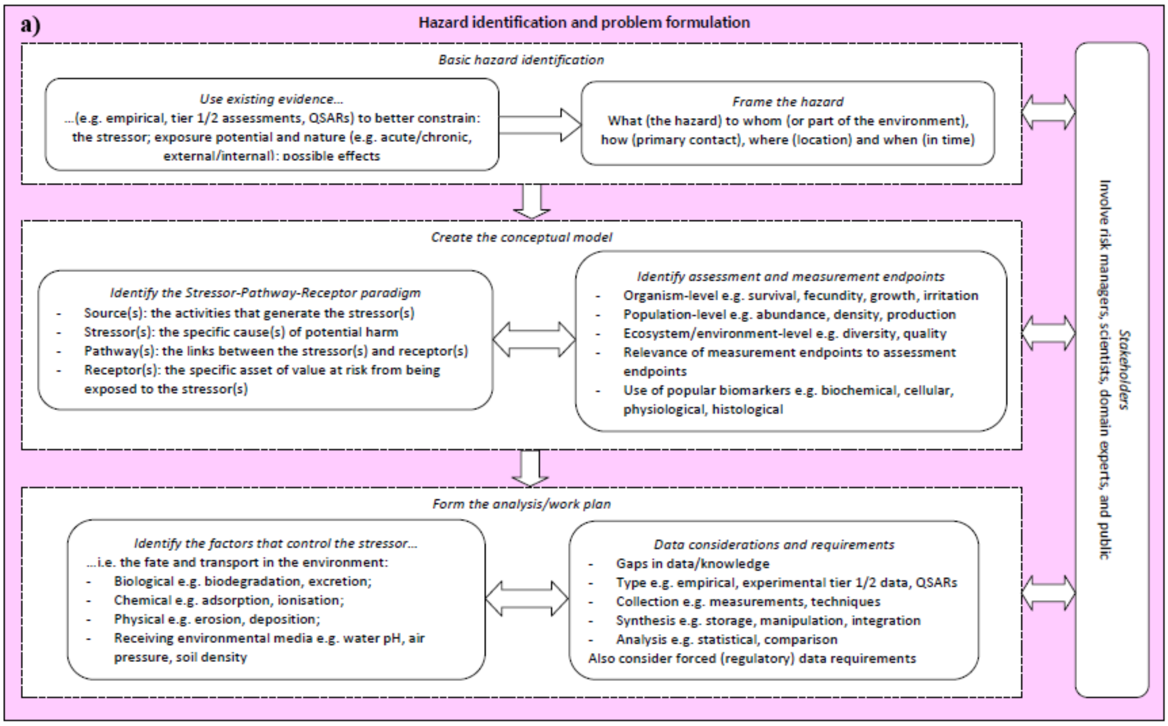


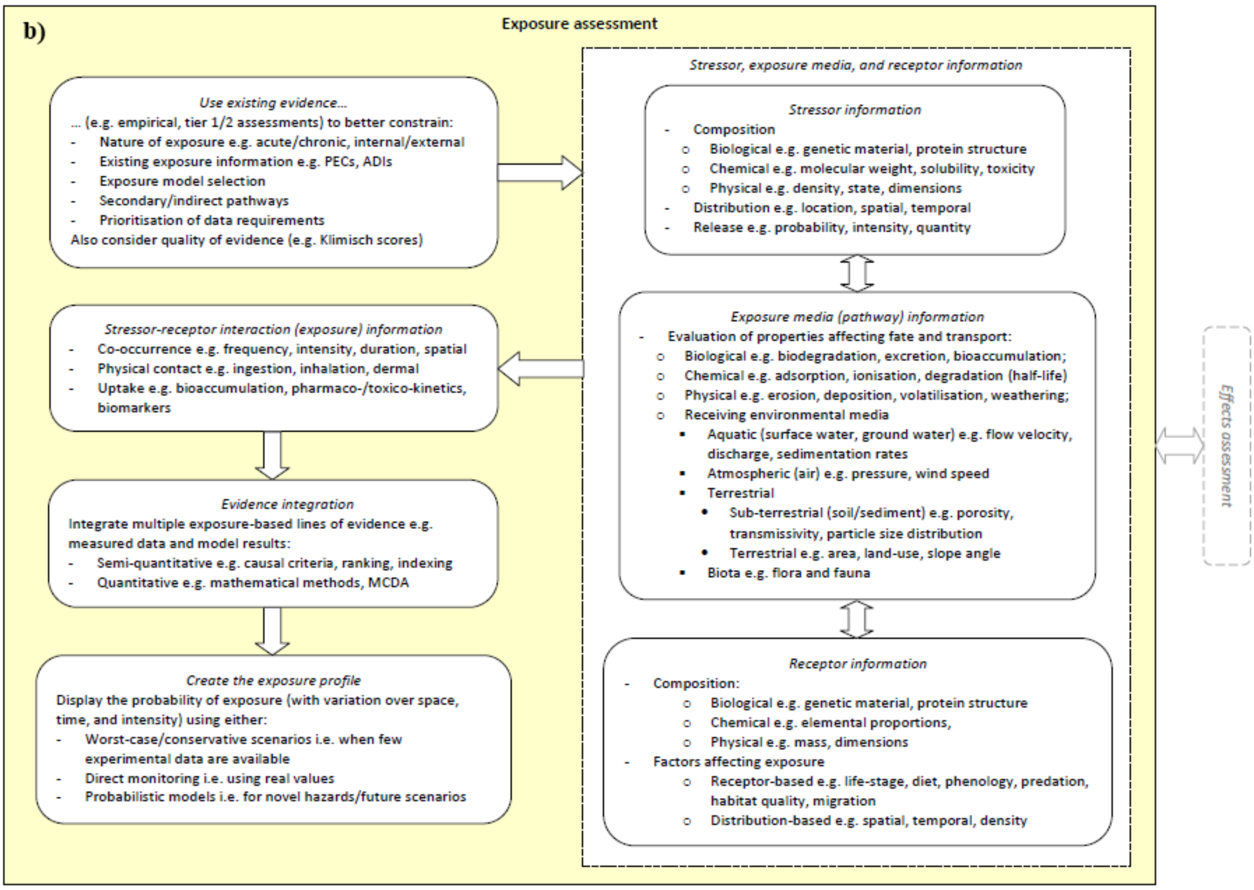


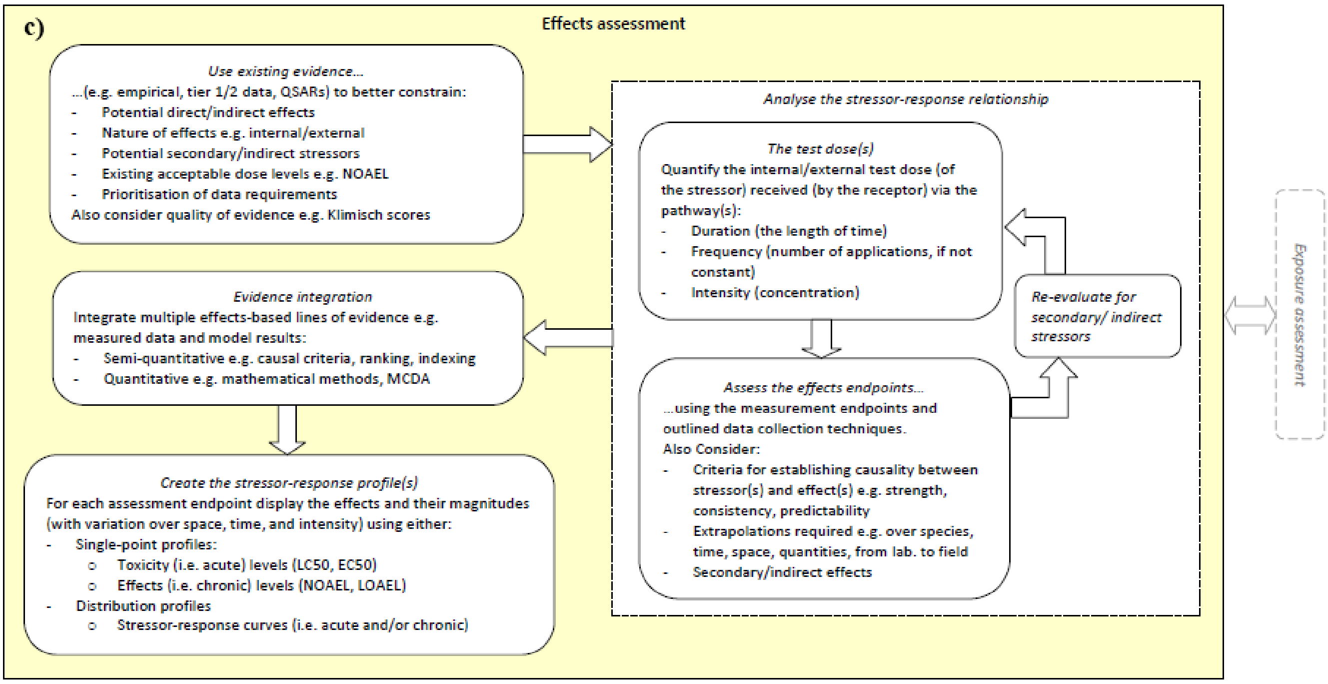


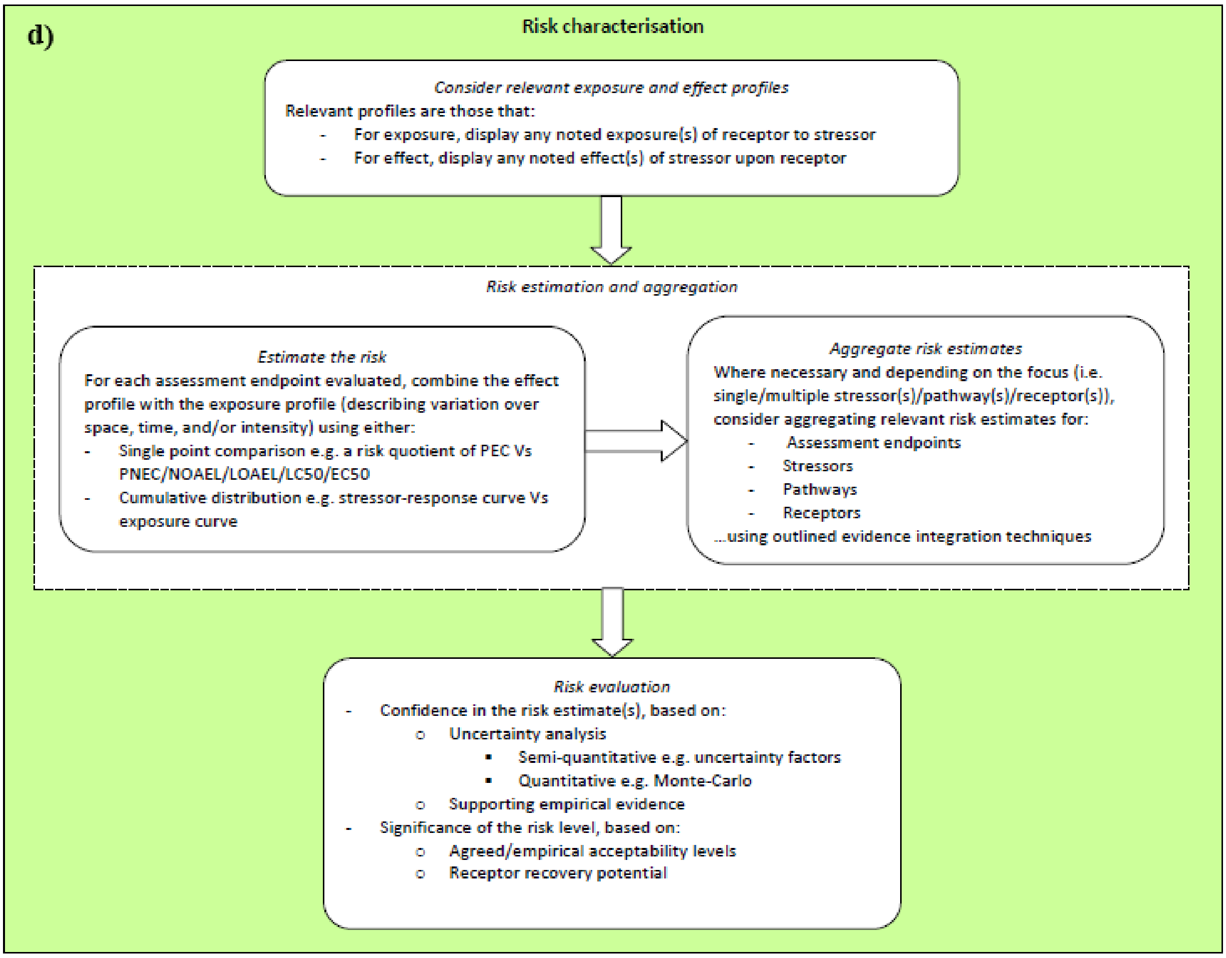
*since this contains no element of probability (i.e. likelihood of consequences being realised) this cannot be considered a true risk characterisation method

Appendix E The generic ERA template, version 2, created through the expert validation of the generic ERA template, version 1, describing the aspects within the stages of: **a)** hazard identification and problem formulation; **b)** exposure assessment; **c)** consequence assessment; and **d)** risk characterisation (see Section 5.3.2).

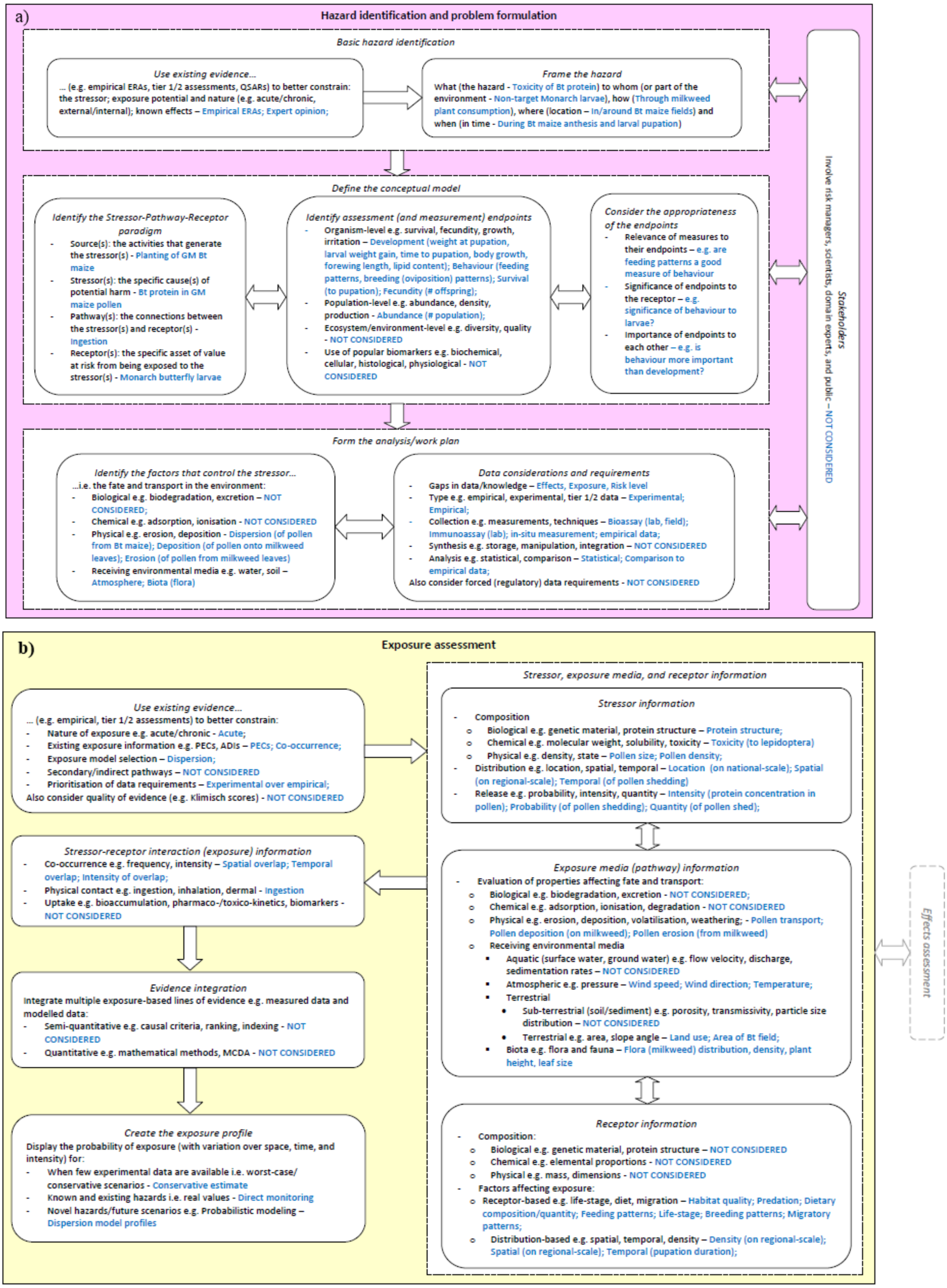





c)
Effects assessment
Use existing evidence...
...(e.g. empirical, tier 1/2 data, QSARs) to better constrain:
- Potential direct/indirect effects
- Nature of effects e.g. internal/external
- Potential secondary/indirect stressors
- Existing acceptable dose levels e.g. NOAEL
- Prioritisation of data requirements
Also consider quality of evidence e.g. Klimisch scores
Evidence integration
Integrate multiple effects-based lines of evidence e.g. measured data and model results:
- Semi-quantitative e.g. causal criteria, ranking, indexing
- Quantitative e.g. mathematical methods, MCDA
Create the stressor-response profile(s)
For each assessment endpoint display the effects and their magnitudes (with variation over space, time, and intensity) using either:
- Single-point profiles:
o Toxicity (i.e. acute) levels (LC50, EC50)
o Effects (i.e. chronic) levels (NOAEL, LOAEL)
- Distribution profiles
o Stressor-response curves (i.e. acute and/or chronic)
Analyse the stressor-response relationship
The test dose(s)
Quantify the internal/external test dose (of the stressor) received (by the receptor) via the pathway(s):
- Duration (the length of time)
- Frequency (number of applications, if not constant)
- Intensity (concentration)
Assess the effects endpoints...
...using the measurement endpoints and outlined data collection techniques.
Also Consider:
- Criteria for establishing causality between stressor(s) and effect(s) e.g. strength, consistency, predictability
- Extrapolations required e.g. over species, time, space, quantities, from lab. to field
- Secondary/indirect effects
Re-evaluate for secondary/ indirect stressors
Exposure assessment


d)
Risk characterisation
Consider relevant exposure and effect profiles
Relevant profiles are those that:
- For exposure, display any noted exposure(s) of receptor to stressor
- For effect, display any noted effect(s) of stressor upon receptor
Risk estimation and aggregation
Estimate the risk
For each assessment endpoint evaluated, combine the effect profile with the exposure profile (describing variation over space, time, and/or intensity) using either:
- Single point comparison e.g. a risk quotient of PEC Vs PNEC/NOAEL/LOAEL/LC50/EC50
- Cumulative distribution e.g. stressor-response curve Vs exposure curve
Aggregate risk estimates
Where necessary and depending on the focus (i.e. single/multiple stressor(s)/pathway(s)/receptor(s)), consider aggregating relevant risk estimates for:
- Assessment endpoints
- Stressors
- Pathways
- Receptors
...using outlined evidence integration techniques
Risk evaluation
- Confidence in the risk estimate(s), based on:
o Uncertainty analysis
▪ Semi-quantitative e.g. uncertainty factors
▪ Quantitative e.g. Monte-Carlo
o Supporting empirical evidence
- Significance of the risk level, based on:
o Agreed/empirical acceptability levels
o Receptor recovery potential

Appendix F The Bt-maize risk to non-target Monarch larvae ERA template, version 1, created by populating the generic ERA template, version 3, with information from relevant articles, describing the aspects within the stages of: **a)** hazard identification and problem formulation; **b)** exposure assessment; **c)** effects assessment; and **d)** risk characterisation (see Section 5.4.2).



c)

Effects assessment

Use existing evidence…

… (e.g. empirical, tier 1/2 assessments) to better constrain:

- Potential direct/indirect effects – (see endpoints);
- Nature of effects e.g. internal/external - external
- Potential secondary/indirect stressors - NOT CONSIDERED
- Existing acceptable dose levels e.g. NOAEL - LOAEC
- Prioritisation of data requirements - Experimental over empirical;
- Also consider quality of evidence e.g. Klimisch scores - NOT CONSIDERED

Evidence integration

Integrate multiple effects-based lines of evidence measured data and modelled data:

- Semi-quantitative e.g. causal criteria, ranking, indexing – NOT CONSIDERED
- Quantitative e.g. mathematical methods/models, MCDA – NOT CONSIDERED

Create the stressor-response profile(s)

For each assessment endpoint display the effects and their magnitudes (with variation over space, time, and intensity) using either:

- Single-point profiles:
  - Toxicity (i.e. acute) levels e.g. LC50 - LC50; EC50;
  - Effects (i.e. chronic) levels e.g. NOAEL, LOAEL– NOT CONSIDERED
- Distribution profiles
  - Stressor-response (i.e. acute and/or chronic) – Distributions; Curves;

Analyse the stressor-response relationship

The test dose(s)

Quantify the internal/external test dose (of the stressor) received (by the receptor) via the pathway(s):

- Duration – of exposure to test dose;
- Frequency (number of applications) – constant;
- Intensity – protein concentration in pollen;

Re-evaluate for secondary/ indirect stressors – NOT CONSIDERED

Assess the effect endpoints…

…using the measurement endpoints:

Assessment endpoint	Measure
development	weight gain (periodic)
development	weight at pupation
development	weight at eclosion
development	time to pupation
development	time to eclosion
development	body growth
development	forewing length
development	lipid content
behaviour	leaf consumption;
behaviour	breeding (oviposition) patterns;
survival	survival to pupation
survival	survival to eclosion
fecundity	number of offspring
abundance	population total

… and outlined data collection techniques – Bioassay (lab, field); Immunoassay (lab); In-situ measurement; Empirical data;

Also consider:

- Criteria for establishing causality between stressor(s) and effect(s) e.g. strength, consistency, predictability – Strength of association; Consistency of association;
- Extrapolations required e.g. within/between species, time, space, quantity, from lab. to field – Time; Space; Quantity; Lab to field;

Exposure assessment

d)

Risk characterisation

Consider relevant exposure and effect profiles

Relevant profiles are those that:

- For exposure, display any noted exposure(s) of receptor to stressor
- For effect, display any noted effect(s) of stressor upon receptor

Risk estimation and aggregation

Estimate the risk

For each assessment endpoint evaluated, *combine* the effect profile with the exposure profile (describing variation over space, time, and/or intensity) using either:

- Single point comparison e.g. a risk quotient - PEC Vs LC50; PEC Vs EC50
- Cumulative distribution - stressor-response curve Vs exposure curve

Aggregate risk estimates

Where necessary and depending on the focus (i.e. single/multiple stressor(s)/pathway(s)/receptor(s)), consider aggregating relevant risk estimates for:

- Assessment endpoints – Development; Behaviour; Survival;
- Stressors – Multiple Bt protein events;
- Pathways – NOT CONSIDERED
- Receptors – NOT CONSIDERED

using previously outlined evidence integration techniques

Risk evaluation

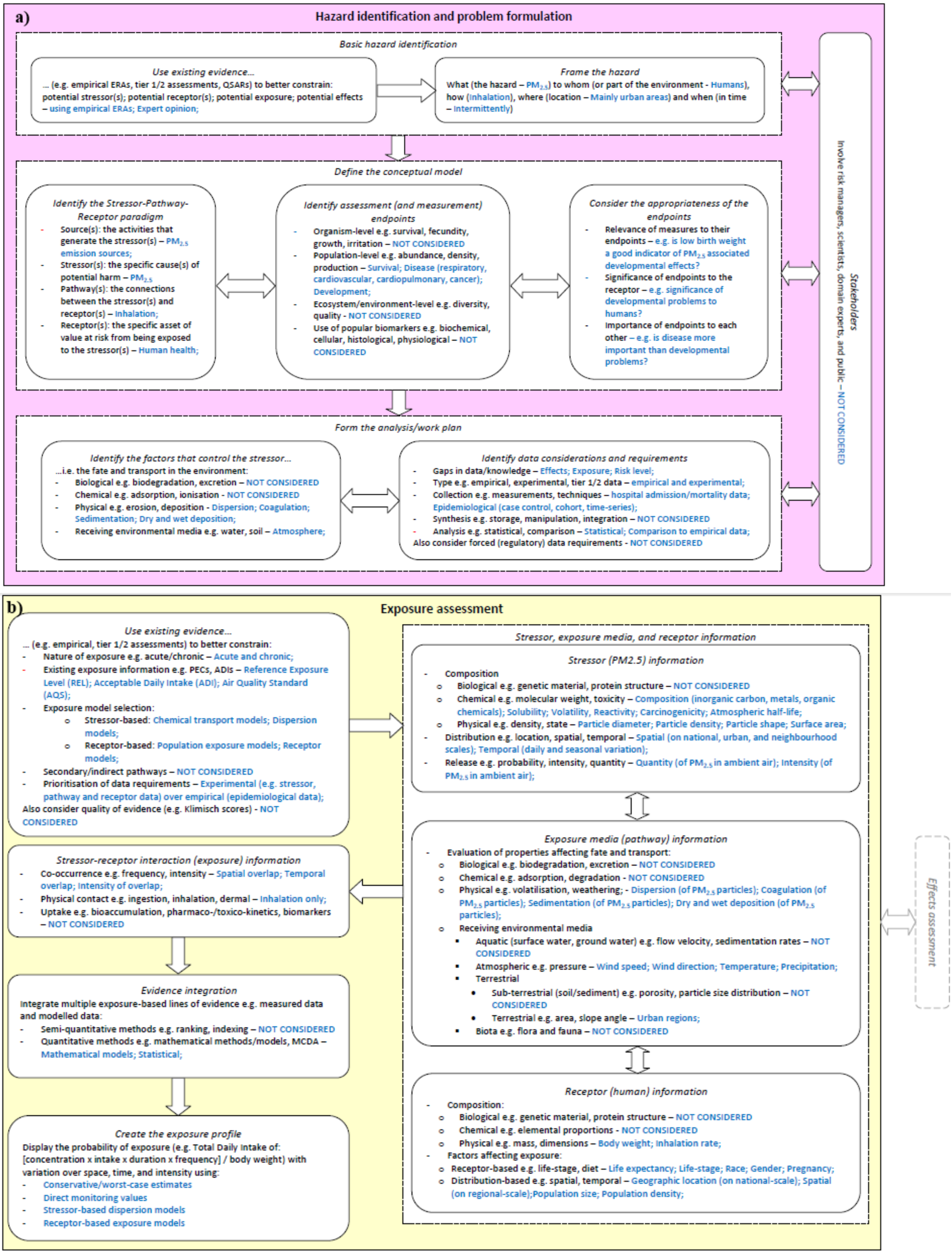
- Confidence in the risk estimate(s), based on:
  - Uncertainty analysis – NOT CONSIDERED
    - Semi-quantitative e.g. uncertainty factors
    - Quantitative e.g. Monte-Carlo
  - Supporting *empirical* evidence – Lab-based; Field-based; Model-based; Expert opinion;
  - Supporting *experimental* evidence – Lab-based; Field-based; Model-based; Expert opinion;
- Significance of the risk level, based on:
  - Regulatory-enforced acceptability levels – i.e. expert-derived and enforced by appropriate regulator(s)
  - Stakeholder-derived acceptability levels – i.e. those involved in current risk assessment
  - Experimentally-derived acceptability levels – i.e. other risk assessments and relevant academic/grey research
  - Receptor recovery potential – NOT CONSIDERED

Appendix G Median occurrence rates (%) for the natures and locations of uncertainty within Case Study 1, organised by ERA task and including ERA phase and overall medians, and the highest proportion(s; of at least 50%) for the nature (shaded blue) and location (shaded red) of uncertainty associated with each ERA task (see Section 5.4.3).

ERA Task	Nature of uncertainty (%)			Location of uncertainty (%)						
	Epistemic	Aleatory	Combined	Data	Language	System	Variability	Extrapolation	Model	Decision
1	20	0	80	80	0	20	60	0	20	20
2	20	0	80	80	0	20	60	60	40	20
3	20	20	60	60	0	60	60	20	40	20
4	20	20	60	60	40	20	60	20	40	20
5	40	20	40	60	40	20	60	20	20	20
6	60	0	20	60	40	20	20	20	20	20
7	20	0	20	40	20	20	20	0	20	20
8	20	20	60	60	40	20	60	20	20	60
9	20	20	60	60	0	20	60	60	20	60
10	20	0	20	20	20	0	0	0	0	20
11	40	0	40	60	40	20	20	20	20	20
12	20	0	20	0	20	20	0	0	0	0
13	60	0	0	0	60	20	0	0	0	0
14	20	20	60	20	20	40	20	60	40	60
15	20	0	80	80	0	40	60	20	40	60
17	20	0	60	60	0	40	20	20	40	60
18	20	0	60	20	20	40	20	20	40	60
19	0	20	80	60	20	40	80	60	40	60
22	60	0	40	60	40	60	20	20	40	40
23	0	20	80	60	0	60	80	80	40	40
24	0	0	100	60	40	60	40	40	60	60
27	20	20	60	80	40	60	60	20	20	40
28	20	20	100	100	0	80	100	20	20	20
29	20	0	60	60	0	20	20	0	40	60
30	20	0	80	40	20	20	0	0	40	80
31	20	0	80	40	20	20	0	0	40	80
32	40	0	60	40	20	20	0	0	40	60
Problem formulation median	**20**	**0**	**60**	**60**	**20**	**20**	**20**	**20**	**40**	**40**
33	60	0	20	60	20	60	0	0	20	0
34	20	0	60	60	0	60	20	20	60	0
35	60	20	0	60	0	20	0	0	20	0
37	60	0	0	40	0	60	0	0	0	0
38	20	0	60	60	0	0	60	0	20	0
39	20	60	20	20	20	0	60	40	20	20
40	0	40	60	60	0	0	80	20	20	0
41	20	20	60	60	0	20	60	20	40	20
42	20	20	60	60	0	20	60	20	40	20
43	20	60	20	40	0	0	80	0	0	0
44	20	60	20	40	0	20	80	20	20	20
45	20	60	20	40	0	20	80	40	20	20
46	20	20	60	80	0	20	80	0	0	0
48	20	0	80	80	0	20	60	20	80	20
49	20	20	60	60	0	0	60	0	20	0

50	20	0	80	80	0	0	60	0	20	0
52	0	60	20	20	0	0	80	20	20	0
55	20	20	60	60	0	20	60	20	60	0
56	0	20	80	20	0	60	80	60	40	0
57	0	20	80	20	0	60	80	60	40	0
58	0	40	60	20	0	20	80	60	60	20
59	20	20	60	60	0	20	60	20	60	20
60	20	20	60	60	0	0	60	60	60	20
65	20	0	80	60	0	60	60	60	60	0
66	20	0	80	40	0	80	60	60	60	0
67	20	0	60	80	20	20	60	60	20	0
68	20	0	60	60	0	20	20	20	60	20
69	20	0	60	60	0	20	0	20	60	20
Exposure assessment median	**20**	**20**	**60**	**60**	**0**	**20**	**60**	**20**	**30**	**0**
70	80	0	20	80	0	0	0	0	20	0
71	80	0	20	60	0	20	0	0	20	0
73	0	40	60	60	0	20	100	80	20	0
74	20	0	60	60	20	20	20	40	40	60
75	20	20	60	60	0	0	60	20	40	0
76	0	60	20	0	0	0	60	20	20	0
77	0	20	80	60	0	20	80	40	40	0
78	20	20	60	60	20	20	60	20	20	0
79	20	0	80	60	20	20	60	20	40	20
81	20	20	60	60	20	20	60	0	20	0
82	20	0	20	20	0	0	0	0	20	0
83	20	20	60	80	20	20	80	40	0	0
88	0	20	80	0	0	0	60	80	80	40
89	0	20	80	0	0	0	60	80	80	40
90	0	20	80	0	0	0	60	80	80	40
91	20	0	80	40	0	0	80	60	60	20
Effects assessment median	**20**	**20**	**60**	**60**	**0**	**10**	**60**	**30**	**30**	**0**
92	20	0	80	40	20	60	60	60	40	60
93	20	20	60	40	20	40	60	60	40	60
94	0	20	80	20	0	20	80	80	80	40
95	0	20	80	20	0	20	80	80	80	40
96	0	20	80	40	0	40	80	60	40	60
97	20	0	60	20	0	60	40	20	40	60
100	20	0	80	80	0	20	20	20	60	60
101	20	20	60	60	0	20	60	60	40	40
102	20	0	80	40	80	40	20	0	20	60
103	40	0	60	40	80	80	0	0	20	60
104	20	0	80	80	20	0	20	20	40	60
Risk characterisation median	**20**	**0**	**80**	**40**	**0**	**40**	**60**	**60**	**40**	**60**
Overall median	**20**	**20**	**60**	**60**	**0**	**20**	**60**	**20**	**40**	**20**

Appendix H The $PM_{2.5}$ risk to human health ERA template, version 1, created by populating the generic ERA template, version 3, with information from relevant articles, describing the aspects within the stages of: **a)** hazard identification and problem formulation; **b)** exposure assessment; **c)** effects assessment; and **d)** risk characterisation (see Section 5.5.2).



c)

Effects assessment (A.K.A. Stressor/Dose/Concentration-response)

Use existing evidence…

… (e.g. empirical, tier 1/2 assessments) to better constrain:

- Potential direct/indirect effects – (see endpoints);
- Nature of effects e.g. internal/external – internal;
- Potential secondary/indirect stressors – NOT CONSIDERED
- Existing toxicity info e.g. LC_{50} – NOT CONSIDERED
- Prioritisation of data requirements - Empirical (epidemiological data) over experimental;

Also consider quality of evidence e.g. Klimisch scores - NOT CONSIDERED

Analyse the stressor-response relationship

The test dose(s)

Quantify the internal/external test dose (of the stressor) received (by the receptor) via the pathway(s):

- Duration – of exposure to test dose (hours/days/weeks);
- Frequency (number of doses) – Intermittent;
- Intensity – $PM_{2.5}$ concentration inhaled;

Re-evaluate for identified secondary/ indirect stressors – NOT CONSIDERED

Assess the effect endpoints…

…using the measurement endpoints:

Assessment endpoint	Measure
Survival (mortality)	Death from PM related disease (cardiovascular/cardiopulmonary, respiratory, cancer)
Disease (cardiovascular/ cardiopulmonary)	Clinical diagnosis
Disease (cancer)	Clinical diagnosis
Disease (respiratory)	Clinical diagnosis
Development	Birth defects, low birth weight, preterm birth

… and outlined data collection techniques – hospital admission/ mortality data; Epidemiological (case control, cohort, time-series);

Also consider:

- Criteria (Bradford-Hill) for establishing causality between stressor(s) and effect(s) e.g. strength, consistency, predictability – Strength of association; Consistency of association;
- Extrapolations required e.g. within/between species, time, space, quantity, from lab. to field – Intraspecies (humans→humans); Quantity (of dose received); Time (duration of exposure to dose);
- Secondary/indirect effects e.g. visibility, climate – NOT CONSIDERED

Evidence integration

Integrate multiple effects-based lines of evidence (e.g. measured data and modelled data) using:

- Semi-quantitative methods e.g. ranking, indexing – NOT CONSIDERED
- Quantitative methods e.g. mathematical methods/models, MCDA – Mathematical models; Statistical;

Create the stressor-response profile(s)

For each assessment endpoint display the effects and their magnitudes (with variation over space, time, and intensity) using either:

- Single-point profiles:
  - Toxicity (i.e. acute) levels e.g. LC_{50} – NOT CONSIDERED
  - Effects (i.e. chronic) levels e.g. NOAEL, LOAEL– Cancer Slope Factor (CSF); Chronic Reference Dose (RfD);
- Distribution profiles
  - Stressor-response (i.e. acute and/or chronic) – Distribution (of intensity of stressor) Vs magnitude of response;

Exposure assessment

d)

Risk characterisation

Consider relevant exposure and effect profiles

Relevant profiles are those that:

- For exposure, display any noted exposure(s) of receptor to stressor
- For effect, display any noted effect(s) of stressor upon receptor

Risk estimation and aggregation

Estimate the risk

For each assessment endpoint evaluated, *combine* the effect profile with the exposure profile (describing variation over space, time, and/or intensity) using either:

- Single point comparison e.g. a risk quotient –
  - Hazard quotient (Exposure ⁄ Effects);
  - Excess life-time risk (Exposure *x* Effects);
- Cumulative distribution – Exposure curve *x* stressor-response curve

Aggregate risk estimates

Where necessary and depending on the focus (i.e. single/multiple stressor(s)/pathway(s)/receptor(s)), consider aggregating relevant risk estimates for:

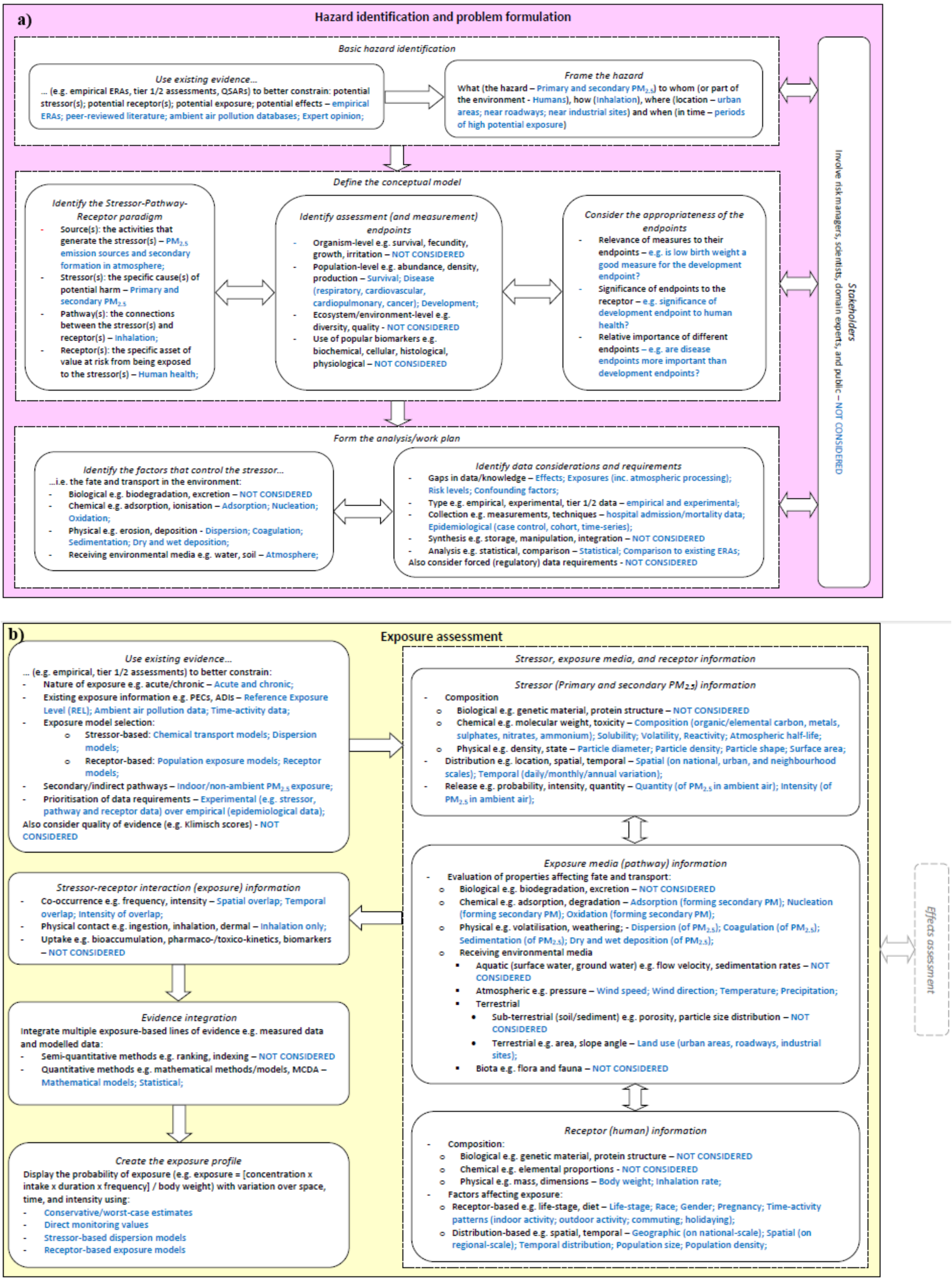
- Assessment endpoints – Development, Disease, and Survival;
- Stressors – Multiple $PM_{2.5}$ events (i.e. of differing composition);
- Pathways – NOT CONSIDERED
- Receptors – NOT CONSIDERED

…using previously outlined evidence integration techniques

Risk evaluation

- Confidence in the risk estimate(s), based on:
  - Uncertainty analysis – NOT CONSIDERED
    - Semi-quantitative e.g. uncertainty factors
    - Quantitative e.g. Monte-Carlo
  - Supporting *empirical* evidence – Epidemiological;
  - Supporting *experimental* evidence – Epidemiological; Toxicological;
- Significance of the risk level, based on:
  - Regulatory-enforced acceptability levels – i.e. expert-derived and enforced by appropriate regulator(s)
  - Stakeholder-derived acceptability levels – i.e. those involved in current risk assessment
  - Experimentally-derived acceptability levels – i.e. other risk assessments and relevant academic/grey research
  - Receptor recovery potential – NOT CONSIDERED

Appendix J The $PM_{2.5}$ risk to human health ERA template, version 2, created through the expert validation of version 1, describing the important aspects within the phases of: **a)** hazard identification and problem formulation; **b)** exposure assessment; **c)** effects assessment; and **d)** risk characterisation (see Section 5.5.2).



c) Effects assessment (A.K.A. Stressor/Dose/Concentration-response)

Use existing evidence…

… (e.g. empirical, tier 1/2 assessments) to better constrain:

- Potential direct/indirect effects – (see endpoints);
- Nature of effects e.g. internal/external – Internal;
- Potential secondary/indirect stressors – Non-ambient $PM_{2.5}$
- Existing toxicity info e.g. LC_{50} – Acceptable Daily Intake (ADI); Air Quality Standard (AQS); Stressor(s) carcinogenicity/toxicity; Animal LC_{50}/EC_{50} values;
- Prioritisation of data requirements - Empirical (epidemiology data) over experimental (epidemiology data);

Also consider quality of evidence e.g. Klimisch scores - NOT CONSIDERED

Analyse the stressor-response relationship

The test dose(s)

Quantify the internal/external test dose (of the stressor) received (by the receptor) via the pathway(s):

- Concentration – of $PM_{2.5}$ inhaled;
- Duration – of exposure to test dose (hours/days/weeks);
- Frequency (number of applications) – of test dose;

Re-evaluate for identified secondary/ indirect stressors – Multiple primary and secondary ambient $PM_{2.5}$ stressors

Assess the effect endpoints…

…using the measurement endpoints:

Assessment endpoint	Measure
Survival (mortality)	Death from PM related disease (cardiovascular/cardiopulmonary, respiratory, cancer)
Incidence of disease (cardiovascular/ cardiopulmonary)	Diagnosis of an associated illness (e.g. myocardial infarction)
Incidence of disease (cancer)	Diagnosis of an associated illness (e.g. malignant tumour)
Incidence of disease (respiratory)	Diagnosis of an associated illness (e.g. reduced pulmonary function)
Development	Birth defects, low birth weight, preterm birth

… and outlined data collection techniques – hospital admission/ mortality data; Epidemiological (case control, cohort, time-series);

Also consider:

- Criteria (e.g. Bradford-Hill) for establishing causality between stressor(s) and effect(s) e.g. strength, predictability – Strength of association; Consistency of association;
- Extrapolations required e.g. within/between species, time, space, quantity, from lab. to field – Intraspecies (humans→humans); Quantity (of dose received); Time (duration of exposure to dose);
- Secondary/indirect effects e.g. visibility, climate – NOT CONSIDERED

Exposure assessment

Evidence integration

Integrate multiple effects-based lines of evidence (e.g. measured data and modelled data) using:

- Semi-quantitative methods e.g. ranking, indexing – NOT CONSIDERED
- Quantitative methods e.g. mathematical methods/models, MCDA – Mathematical models; Statistical;

Create the stressor-response profile(s)

For each assessment endpoint display the effects and their magnitudes (with variation over space, time, and intensity) using either:

- Single-point profiles:
  - Toxicity (i.e. acute) levels e.g. LC_{50} – NOT CONSIDERED
  - Effects (i.e. chronic) levels e.g. NOAEL, LOAEL– Cancer Slope Factor (CSF); Chronic Reference Dose (RfD);
- Distribution profiles
  - Stressor-response (i.e. acute and/or chronic) – Distribution (of intensity of stressor) Vs. magnitude of response;

d) Risk characterisation

Consider relevant exposure and effect profiles

Relevant profiles are those that:

- For exposure, display any exposure(s) of receptor to stressor
- For effect, display any effect(s) of stressor upon receptor

Risk estimation and aggregation

Estimate the risk

For each assessment endpoint evaluated, *combine* the effect profile with the exposure profile (describing variation over space, time, and/or intensity) using either:

- Single point comparison –
  - Risk quotient (Exposure ⁄ Effects);
  - Risk ratio (Pexposed ⁄ Pnon-exposed);
  - Attributable risk (Pexposed–Pnon-exposed);
  - Excess life-time risk (Exposure x Effects);
  - Adjusted life-time risk (DALYs, QALYs);
- Cumulative distribution – Exposure curve Vs. stressor-response curve

Aggregate risk estimates

Where necessary and depending on the focus (i.e. single/multiple stressor(s)/pathway(s)/receptor(s)), consider aggregating relevant risk estimates for:

- Assessment endpoints – Development, Disease, and Survival;
- Stressors – Multiple primary and secondary $PM_{2.5}$ stressors;
- Pathways – NOT CONSIDERED
- Receptors – NOT CONSIDERED

…using previously outlined evidence integration techniques

Risk evaluation

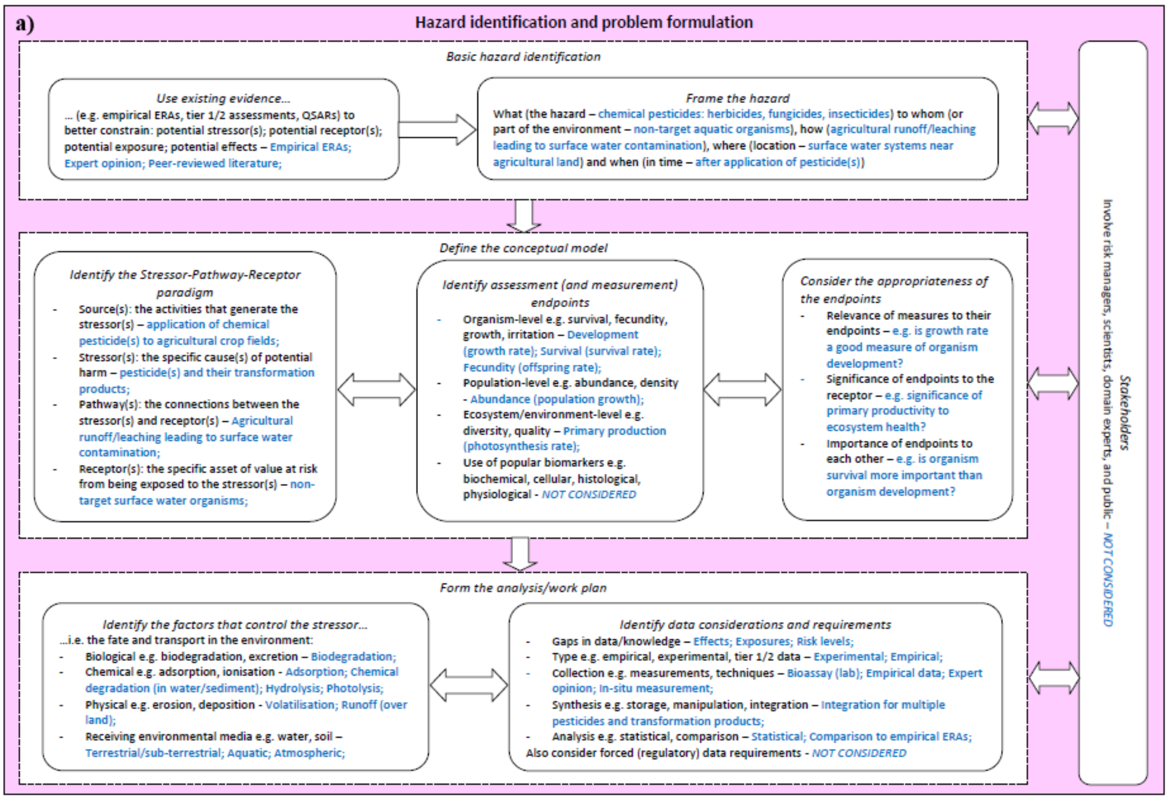
- Confidence in the risk estimate(s), based on:
  - Uncertainty analysis – NOT CONSIDERED *(in the context of this template validation exercise, at least)*
    - Semi-quantitative e.g. uncertainty factors
    - Quantitative e.g. Monte-Carlo
  - Supporting *empirical* evidence – Epidemiological;
  - Supporting *experimental* evidence – Epidemiological; Toxicological;
- Significance of the risk level, based on:
  - Regulatory-enforced acceptability levels – i.e. expert-derived and enforced by appropriate regulator(s)
  - Stakeholder-derived acceptability levels – i.e. those involved in current risk assessment
  - Experimentally-derived acceptability levels – i.e. other risk assessments and relevant academic/grey research
  - Receptor recovery potential – NOT CONSIDERED

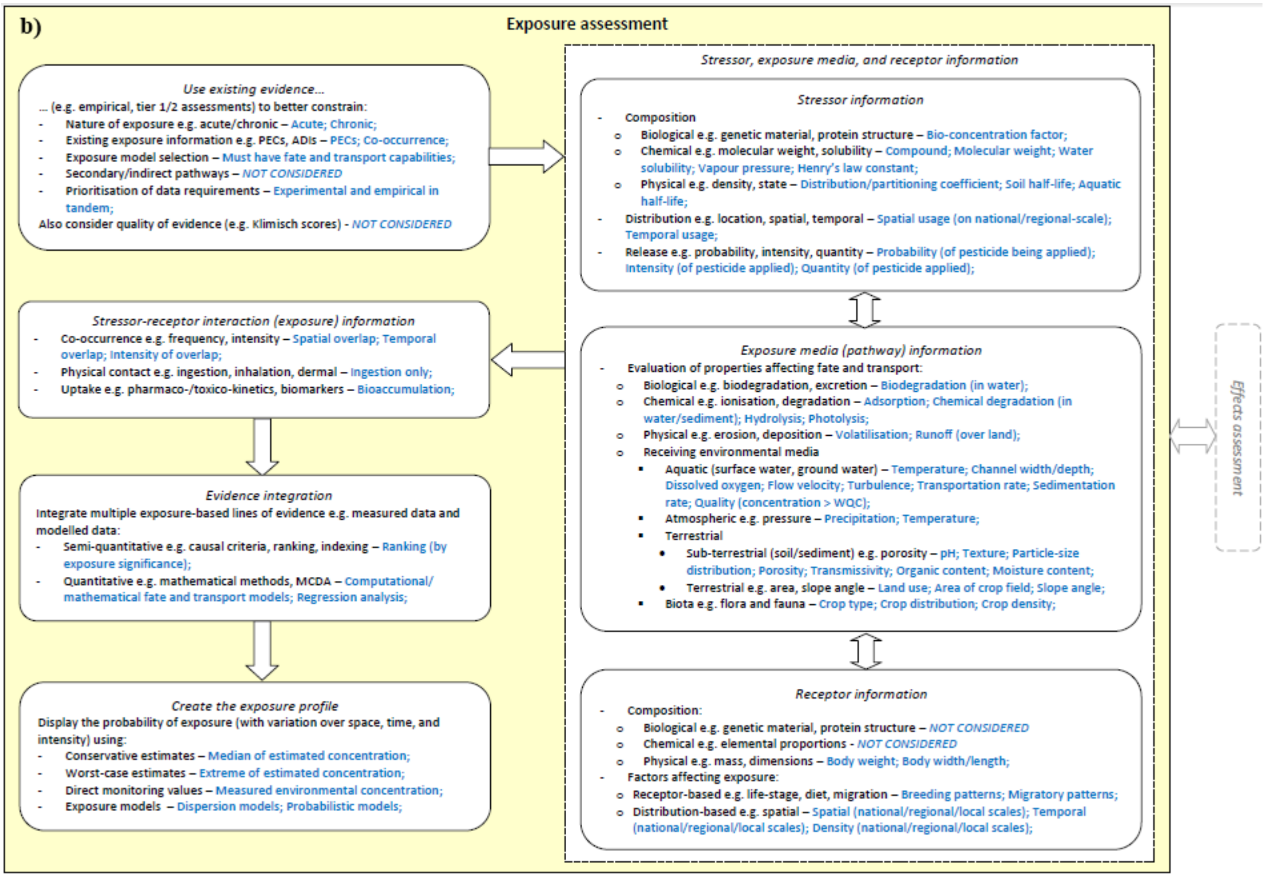
Appendix K Median occurrence rates (%) for the natures and locations of uncertainty within Case Study 2, organised by ERA task and including ERA phase and overall medians, and the highest proportion(s; of at least 50%) for the nature (shaded blue) and location (shaded red) of uncertainty associated with each ERA task (see Section 5.5.3).

ERA Task	Nature of uncertainty (%)			Location of uncertainty (%)						
	Epistemic	**Aleatory**	**Combined**	**Data**	**Language**	**System**	**Variability**	**Extrapolation**	**Model**	**Decision**
1	0	0	100	80	20	100	80	80	80	80
2	20	0	80	80	20	100	80	0	80	60
3	20	0	80	80	20	80	80	20	80	20
4	20	0	80	100	20	80	80	20	80	60
5	20	0	80	80	40	80	40	20	80	0
6	20	0	60	60	20	60	20	0	60	0
7	0	20	40	0	40	40	0	40	40	0
8	20	0	80	80	40	80	40	25	80	25
9	20	0	80	80	80	80	40	20	80	20
10	20	0	80	100	20	100	80	20	40	60
11	20	0	80	100	20	100	80	20	80	20
12	40	0	40	40	20	80	40	0	40	20
13	20	0	80	80	20	80	20	0	40	60
14	0	0	100	80	20	100	80	60	100	0
16	0	0	100	60	20	100	80	60	80	0
17	0	0	100	40	0	100	40	60	80	40
22	0	0	100	80	0	100	80	20	80	20
23	0	20	80	60	0	80	80	60	60	60
24	0	0	100	100	20	100	100	60	100	60
26	20	0	80	100	20	80	80	0	40	20
27	20	0	80	100	20	80	80	0	40	20
28	40	0	60	100	0	80	40	20	40	0
29	0	0	100	100	40	80	80	80	60	100
30	0	0	100	100	20	40	80	80	100	20
31	60	0	40	100	20	20	40	20	20	40
32	40	20	40	80	20	20	40	40	20	40
Problem formulation median	**20**	**0**	**80**	**80**	**20**	**80**	**80**	**20**	**80**	**20**
33	0	0	100	80	40	80	80	40	80	80
34	0	0	100	100	20	80	100	40	80	80
35	0	0	100	100	40	80	80	40	100	80
36	0	20	80	80	20	40	100	40	80	40
37	0	40	20	20	0	20	20	60	20	20
39	20	0	80	100	20	80	80	40	80	20
40	20	0	80	100	20	80	80	40	80	0
41	0	20	80	80	20	40	100	80	80	20
42	0	20	80	80	20	40	100	80	80	20
43	0	0	100	100	20	40	100	40	20	20
45	0	0	80	80	20	40	80	40	40	20
47	0	20	80	80	20	80	100	80	80	20
48	0	0	100	100	20	80	100	80	80	40
49	40	20	40	80	0	20	40	20	20	20
52	20	0	80	100	20	40	80	20	40	40
54	0	0	80	80	0	40	80	60	80	0

55	0	40	40	20	0	40	20	60	40	0
56	0	60	40	40	0	40	80	80	40	20
57	0	20	80	80	0	40	100	80	80	0
58	0	60	40	40	0	20	100	20	40	20
59	0	60	40	40	20	20	100	20	40	0
60	20	20	60	80	0	20	100	60	80	0
61	60	0	40	40	0	100	20	20	40	20
64	20	0	80	80	60	100	80	60	80	80
65	20	20	60	40	0	60	60	40	20	40
66	40	0	60	60	0	40	40	20	40	20
67	20	20	60	60	0	40	80	20	40	0
68	0	0	80	60	0	40	80	20	80	20
69	0	0	100	60	0	80	80	60	100	20
Exposure assessment median	**0**	**0**	**80**	**80**	**20**	**40**	**80**	**40**	**80**	**20**
70	0	0	80	80	20	80	80	40	80	20
71	0	0	80	80	20	80	80	40	80	40
72	0	0	80	80	20	80	80	80	80	20
73	0	0	80	80	20	80	80	80	80	40
74	0	0	100	100	20	40	80	100	80	80
75	0	0	100	100	20	80	100	60	100	40
76	0	20	80	80	20	40	100	60	80	20
77	0	20	80	80	20	40	100	60	80	20
78	0	0	100	100	0	80	100	60	80	60
80	0	0	100	100	0	80	100	80	80	60
81	0	20	80	80	0	20	100	80	80	20
87	0	0	100	100	20	80	80	80	100	60
88	0	0	100	100	0	40	100	60	80	20
89	0	0	100	100	0	40	100	60	80	20
90	0	0	100	100	0	40	100	60	80	20
91	0	0	100	100	0	40	100	60	80	20
Effects assessment median	**0**	**0**	**100**	**100**	**20**	**60**	**100**	**60**	**80**	**20**
92	20	0	80	100	0	80	80	60	40	0
93	20	0	80	100	0	80	80	60	80	20
94	0	0	100	100	40	40	100	80	80	80
95	0	0	100	100	80	40	100	100	100	40
96	0	0	100	80	20	40	80	80	80	60
97	0	0	80	80	20	40	80	60	80	20
100	0	20	80	80	20	40	80	80	20	20
101	0	0	100	80	40	40	80	80	80	20
102	0	20	80	80	20	40	80	80	40	60
103	0	0	100	80	20	40	80	60	40	40
104	0	0	100	80	20	40	80	60	40	40
Risk characterisation median	**0**	**0**	**100**	**80**	**20**	**40**	**80**	**80**	**80**	**40**
Overall median	**0**	**0**	**80**	**80**	**20**	**70**	**80**	**60**	**80**	**20**

Appendix L The agricultural chemical pesticides risk to surface water organisms ERA template, version 1, created by populating the generic ERA template, version 3, with information from relevant articles, describing the aspects within the stages of: **a)** hazard identification and problem formulation; **b)** exposure assessment; **c)** effects assessment; and **d)** risk characterisation (see Section 5.6.2).





c) **Effects assessment (A.K.A. Stressor/Dose/Concentration-response)**

Use existing evidence…

… (e.g. empirical, tier 1/2 assessments) to better constrain:

- Potential direct/indirect effects – (see endpoints);
- Nature of effects e.g. internal/external – External;
- Potential secondary/indirect stressors – Multiple pesticides;
- Existing toxicity levels e.g. NOAEL – Water Quality Criteria (WQC); LC_{50}; EC_{50}; NOAEL; PNEC;
- Prioritisation of data requirements - Experimental and empirical in tandem;

Also consider quality of evidence e.g. Klimisch scores - *NOT CONSIDERED*

Analyse the stressor-response relationship

The test dose(s)

Quantify the internal/external test dose (of the stressor) received (by each receptor) via the pathway(s):

- Concentration – of test dose;
- Duration – of exposure to test dose;
- Frequency (number of applications) – of test dose;
- Quantity – of pesticide(s) in freshwater sample;

Re-evaluate for secondary/ indirect stressors – Multiple pesticides and their transformation products;

Assess the effect endpoints…

…using the measurement endpoints:

Assessment endpoint	Measure
Organism development	Growth rate
Organism survival	Survival rate
Organism fecundity	Offspring rate
Population abundance	Population growth
Ecosystem primary production	Photosynthesis rate

… and outlined data collection techniques – Bioassay (lab); Empirical data; Expert opinion; In-situ measurement;

Also consider:

- Criteria for establishing causality between stressor(s) and effect(s) e.g. strength, consistency, predictability – Strength of association; Consistency of association;
- Extrapolations required e.g. time, space – Within/between species; Time; Space; Quantity; Lab to field;
- Secondary/indirect effects – Changes in water quality; Reductions in herbivore populations;

Evidence integration

Integrate multiple effects-based lines of evidence e.g. measured data and modelled data:

- Semi-quantitative e.g. causal criteria, ranking, indexing – Ranking (by effects significance);
- Quantitative e.g. mathematical methods/models, MCDA – Mathematical models (e.g. concentration addition); Regression analysis;

Create the stressor-response profile(s)

For each assessment endpoint display the effects and their magnitudes (with variation over space, time, and intensity) using either:

- Single-point profiles:
  - Toxicity (i.e. acute) levels e.g. LC50 - LC_{50}; EC_{50};
  - Effects (i.e. chronic) levels e.g. NOAEL, LOAEL– NOAEL; PNEC (NOAEL / assessment factor);
- Distribution profiles
  - Stressor-response (i.e. acute and/or chronic) – Distribution (of intensity of stressor Vs. magnitude of response);

Exposure assessment

d) **Risk characterisation**

Consider relevant exposure and effect profiles

Relevant profiles are those that:

- For exposure, display any exposure(s) of receptor to stressor;
- For effect, display any effect(s) of stressor upon receptor;

Risk estimation and aggregation

Estimate the risk

For each assessment endpoint evaluated, *combine* the effect profile with the exposure profile (describing variation over space, time, and/or intensity) using either:

- Single point comparison:
  - Risk quotient for individual pesticide: PEC / LC_{50}; PEC / EC_{50}; PEC / NOAEL; PEC / PNEC;
  - Risk quotient for multiple pesticides: Sum of individual risk quotients;
- Cumulative/Frequency distribution:
  - For individual/multiple pesticide(s): Exposure curve Vs. stressor-response curve;

Aggregate risk estimates

Where necessary and depending on the focus (i.e. single/multiple stressor(s)/pathway(s)/receptor(s)), consider aggregating relevant risk estimates for:

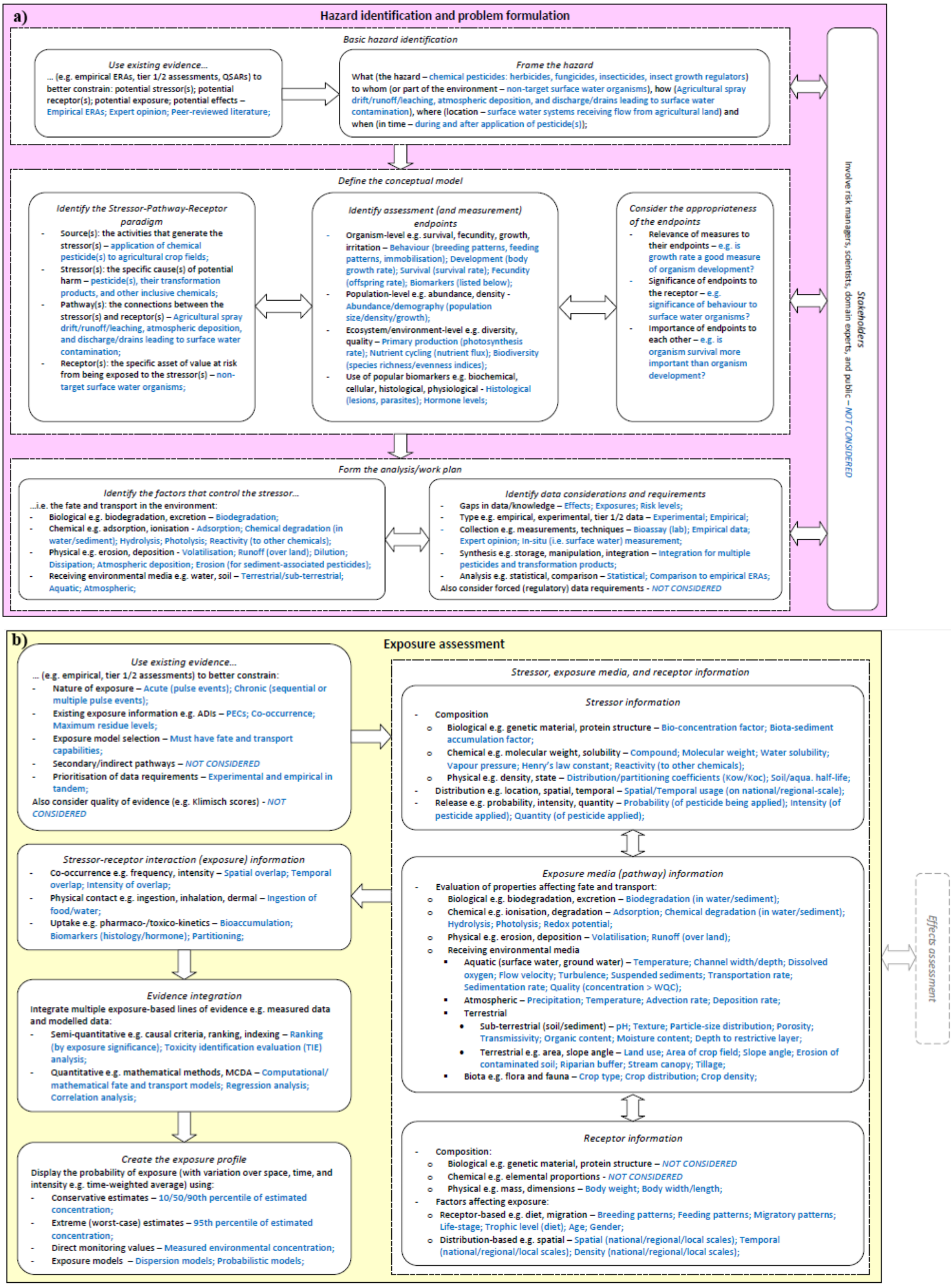
- Assessment endpoints – Development; Survival; Fecundity; Abundance; Primary production;
- Stressors – Multiple pesticides; Multiple transformation products;
- Pathways – *NOT CONSIDERED*
- Receptors – Multiple freshwater organisms e.g. fish, algae, microcrustacea;

…using previously outlined evidence integration techniques.

Risk evaluation

- Confidence in the risk estimate(s), based on:
  - Uncertainty analysis – *NOT CONSIDERED (in the context of this template validation exercise, at least)*
    - Semi-quantitative e.g. Assessment factors
    - Quantitative e.g. Monte-Carlo
  - Supporting empirical (i.e. peer-reviewed) evidence – e.g. Lab-based; Field-based; Model-based; Expert opinion;
  - Supporting experimental (i.e. non peer-reviewed) evidence – e.g. Lab-based; Field-based; Model-based; Expert opinion;
- Significance of the risk level, based on:
  - Regulatory-enforced acceptability levels – i.e. Expert-derived and enforced by appropriate regulator(s);
  - Stakeholder-derived acceptability levels – i.e. Those involved in current risk assessment;
  - Experimentally-derived acceptability levels – i.e. Other risk assessments and relevant academic/grey research;
  - Receptor recovery potential – Recovery from acute low-level exposure;

Appendix M The agricultural chemical pesticides risk to surface water organisms ERA template, version 2, created through the expert validation of version 1, describing the important aspects within the phases of: **a)** hazard identification and problem formulation; **b)** exposure assessment; **c)** effects assessment; and **d)** risk characterisation (see Section 5.6.2).



c) **Effects assessment (A.K.A. Stressor/Dose/Concentration-response)**

Use existing evidence…

… (e.g. empirical, tier 1/2 assessments) to better constrain:
- Potential direct/indirect/secondary effects – (see endpoints);
- Nature of effects e.g. internal/external – Internal; External;
- Potential secondary/indirect stressors – Multiple pesticides;
- Existing toxicity levels – $LC_{10/50}$; EC_{50}; NOAEL; PNEC; Species sensitivity distributions (SSDs); Water Quality Criteria (WQC); EPA aquatic life benchmarks;
- Prioritisation of data requirements - Experimental and empirical in tandem;

Also consider quality of evidence e.g. Klimisch scores - *NOT CONSIDERED*

Analyse the stressor-response relationship

The test dose(s)

Quantify the internal/external test dose (of the stressor) received (by each receptor) via the pathway(s):
- Concentration – of test dose;
- Duration – of exposure to test dose;
- Frequency (number of applications) – of test dose;
- Quantity – of pesticide(s) in freshwater sample;

Re-evaluate for multiple stressors – Multiple pesticides and their transformation products;

Assess the effect endpoints…

…using the measurement endpoints:

Assessment endpoint	Measure
Organism behaviour	Breeding patterns; feeding patterns; immobilisation;
Organism development	Body growth rate
Organism survival	Survival rate
Organism fecundity	Offspring rate
Organism biomarkers	Histology (lesions, parasites); hormone levels
Population abundance/demography	Population size/density/growth
Ecosystem primary production	Photosynthesis rate
Ecosystem nutrient cycling	Nutrient flux
Ecosystem biodiversity	Species richness/evenness indices

… and outlined data collection techniques – Bioassay (lab); Empirical data; Expert opinion; In-situ measurement;

Also consider:
- Criteria for establishing causality between stressor(s) and effect(s) e.g. strength, consistency, predictability – Strength of association; Consistency of association;
- Extrapolations required e.g. time, space – Within/between species; Time; Space; Quantity; Lab to field;
- Secondary/indirect effects – Changes in water quality (WQC); Reductions in aquatic/terrestrial herbivore populations;

Exposure assessment

Evidence integration

Integrate multiple effects-based lines of evidence e.g. measured data and modelled data:
- Semi-quantitative e.g. causal criteria, ranking, indexing – Ranking (by effects significance);
- Quantitative e.g. mathematical methods/models, MCDA – Mathematical models (e.g. concentration addition); Regression analysis; Correlation analysis; Spatially explicit models;

Create the stressor-response profile(s)

For each assessment endpoint display the effects and their magnitudes (with variation over space, time, and intensity) using either:
- Single-point profiles:
  - Conservative toxicity (i.e. acute) levels - LC_{50}; EC_{50};
  - Extreme toxicity (i.e. acute) levels - LC_{10};
  - Effects (i.e. chronic) levels e.g. NOAEL, LOAEL– NOAEL; PNEC (NOAEL / assessment factor);
- Distribution profiles
  - Species sensitivity (SSD) - lognormal of toxicity or effects profiles;
  - Stressor-response (i.e. acute and/or chronic) –distribution (of intensity of stressor Vs. magnitude of response);

d) **Risk characterisation**

Consider relevant exposure and effect profiles

Relevant profiles are those that:
- For exposure, display any exposure(s) of receptor to stressor;
- For effect, display any effect(s) of stressor upon receptor;

Risk estimation and aggregation

Estimate the risk

For each assessment endpoint evaluated, *combine* the effect profile with the exposure profile (describing variation over space, time, and/or intensity) using either:
- Single point comparison:
  - Risk quotient for individual pesticide: PEC / LC_{50}; PEC / EC_{50}; PEC / NOAEL; PEC / PNEC;
  - Risk quotient for multiple pesticides: Sum of individual risk quotients;
  - Quality criterion for individual/multiple pesticide(s): HC_5 value of a SSD;
- Cumulative/Frequency distribution:
  - Risk for individual/multiple pesticide(s):
    - Expected ecological risk: Exposure curve Vs. stressor-response curve;
    - Potentially affected fraction (PAF) = SSD + PEC;

Aggregate risk estimates

Where necessary and depending on the focus (i.e. single/multiple stressor(s)/pathway(s)/receptor(s)), consider aggregating relevant risk estimates for:
- Assessment endpoints – Behaviour; Development; Survival; Fecundity; Abundance/demography; Primary production; Nutrient cycling; Biodiversity;
- Stressors – Multiple pesticides; Multiple transformation products;
- Pathways – Spray drift; Runoff; Leaching; Atmospheric deposition; Discharge/drains;
- Receptors – Multiple freshwater organisms e.g. fish, algae, microcrustacea;

…using previously outlined evidence integration techniques.

Risk evaluation

- Confidence in the risk estimate(s), based on:
  - Uncertainty analysis – *NOT CONSIDERED (in the context of this template validation exercise, at least)*
    - Semi-quantitative e.g. Assessment factors
    - Quantitative e.g. Monte-Carlo
  - Supporting empirical (i.e. peer-reviewed) evidence – e.g. Lab-based; Field-based; Model-based; Expert opinion;
  - Supporting experimental (i.e. non peer-reviewed) evidence – e.g. Lab-based; Field-based; Model-based; Expert opinion;
- Significance of the risk level, based on:
  - Regulatory-enforced acceptability levels – i.e. Expert-derived and enforced by appropriate regulator(s);
  - Stakeholder-derived acceptability levels – i.e. Those involved in current risk assessment;
  - Experimentally-derived acceptability levels – i.e. Other risk assessments and relevant academic/grey research;
  - Receptor recovery potential – Recovery from acute low-level exposure;

Appendix N Median occurrence rates (%) for the natures and locations of uncertainty within Case Study 3, organised by ERA task and including ERA phase and overall medians, and the highest proportion(s; of at least 50%) for the nature (shaded blue) and location (shaded red) of uncertainty associated with each ERA task (see Section 5.6.3).

ERA Task	**Nature of uncertainty (%)**			**Location of uncertainty (%)**						
	Epistemic	**Aleatory**	**Combined**	**Data**	**Language**	**System**	**Variability**	**Extrapolation**	**Model**	**Decision**
1	11	11	67	78	0	67	56	56	44	11
2	33	0	67	89	0	67	67	67	33	11
3	0	0	100	67	11	78	78	67	56	22
4	0	0	100	78	0	56	89	78	67	11
5	44	0	33	33	33	33	11	33	33	11
6	33	11	56	56	33	56	44	44	22	11
7	11	0	89	67	22	44	67	78	78	22
8	22	0	78	78	22	67	44	56	67	11
9	22	0	78	67	0	56	56	67	56	11
10	33	22	44	22	0	33	33	33	22	11
11	33	0	67	44	0	78	33	44	56	11
12	22	11	67	56	0	44	33	22	67	11
13	11	11	78	67	0	56	78	56	44	11
14	33	11	56	56	11	44	44	44	22	0
15	22	11	67	67	11	33	44	56	33	0
17	22	33	44	33	11	33	56	44	22	0
18	22	33	44	44	11	33	56	44	22	0
19	22	11	67	67	22	44	56	56	56	33
20	22	0	78	89	33	67	56	56	67	33
21	22	0	78	67	44	89	56	56	67	33
22	44	11	44	67	11	44	33	33	11	0
23	22	0	78	33	0	67	78	67	56	11
24	11	11	78	33	33	67	56	78	67	44
25	22	11	67	67	11	67	56	44	56	22
26	11	33	56	44	0	33	78	56	44	11
27	11	11	78	67	0	67	78	67	44	11
28	22	22	56	44	0	44	67	56	44	11
29	33	0	67	56	33	78	11	22	56	22
30	44	0	56	56	33	44	11	22	33	22
31	44	11	33	56	22	33	22	22	11	11
32	44	0	56	67	33	33	22	33	44	22
Problem formulation median	**22**	**11**	**67**	**67**	**11**	**56**	**56**	**56**	**44**	**11**
33	22	0	67	67	0	67	56	56	56	11
34	44	11	44	56	11	56	33	11	11	11
35	22	33	44	56	22	56	56	56	33	11
37	0	11	89	56	11	67	56	67	89	33
38	22	0	78	56	33	89	44	44	44	33
39	33	33	22	44	0	11	56	33	11	0
40	44	22	22	67	11	0	33	11	11	0
41	33	22	33	56	0	22	56	22	11	22
42	33	0	67	100	0	33	56	44	44	11

43	33	0	67	100	0	44	67	44	33	22
44	33	0	67	89	0	44	44	44	22	22
45	33	11	56	89	0	22	56	33	22	11
46	44	0	56	100	0	33	44	33	33	22
47	33	11	56	67	0	44	56	44	44	11
48	44	22	33	67	11	33	44	44	22	11
49	33	22	44	56	0	44	44	44	33	11
50	33	11	44	56	11	22	44	22	22	11
51	44	11	44	89	0	11	56	11	11	11
52	44	11	33	67	0	22	44	11	11	0
53	11	22	56	56	0	0	56	22	22	0
54	11	11	67	56	0	33	67	44	22	0
55	33	33	33	67	0	0	44	33	0	0
56	44	22	33	78	0	11	33	22	0	0
57	33	33	33	67	0	44	33	56	0	11
58	22	22	56	78	0	44	56	56	11	11
59	22	22	56	67	0	44	56	33	22	11
60	22	11	67	67	0	56	44	22	33	11
61	22	11	67	56	0	33	67	22	44	11
62	44	11	44	67	0	33	33	11	22	11
63	11	11	78	67	0	33	67	11	44	11
64	11	0	78	56	22	56	56	56	67	22
65	22	0	67	67	11	78	56	44	67	22
66	22	33	44	44	11	11	33	56	11	22
67	33	33	33	56	11	11	33	56	22	22
68	11	33	56	67	0	22	67	11	22	33
69	11	0	89	33	0	67	44	56	78	22
Exposure assessment median	**0**	**0**	**100**	**33**	**0**	**56**	**22**	**67**	**89**	**33**
70	33	11	44	56	0	33	56	44	22	11
71	56	0	33	67	22	67	22	22	44	0
72	44	11	44	67	0	56	22	44	33	0
73	33	0	67	67	22	56	33	33	56	11
74	11	11	78	67	0	56	44	67	22	22
75	22	11	67	33	33	11	22	33	33	22
76	33	0	56	67	0	22	22	33	44	0
77	22	11	67	56	0	44	44	56	56	0
78	33	11	56	78	0	22	56	22	44	0
79	11	22	67	67	0	11	67	11	11	0
81	11	22	67	67	0	33	67	22	11	0
82	11	11	67	56	0	11	56	33	11	0
83	11	22	67	56	0	22	67	22	11	0
84	0	11	89	67	0	78	67	67	56	11
85	11	0	89	89	22	78	67	67	67	11
86	11	0	89	78	22	89	67	78	56	22
87	22	11	67	56	0	56	56	33	56	11
88	0	0	100	67	0	56	67	56	100	22
89	33	0	67	89	0	33	33	33	44	0
90	22	0	67	78	0	33	33	33	44	0
91	22	11	67	78	0	44	44	67	56	0
Effects assessment median	**11**	**0**	**89**	**78**	**0**	**33**	**67**	**67**	**78**	**0**
92	11	0	89	67	0	33	67	33	78	33
93	11	11	78	67	0	44	67	78	78	22
94	0	11	89	44	0	56	44	78	78	33
95	0	11	89	33	11	56	67	78	67	11
96	0	0	100	33	0	67	44	67	78	33
97	0	0	100	33	0	78	44	89	78	33

98	0	0	100	33	0	78	44	67	78	33
99	0	0	100	44	0	78	67	89	78	33
100	0	11	89	44	0	44	67	89	44	22
101	0	11	89	44	0	44	44	89	33	22
102	11	11	78	33	33	56	44	67	44	33
103	0	22	78	22	22	56	44	78	56	44
104	0	11	89	33	22	44	67	78	78	33
105	11	0	89	67	11	78	67	78	89	33
Risk characterisation median	**0**	**11**	**89**	**39**	**0**	**56**	**56**	**78**	**78**	**33**
Overall median	**22**	**11**	**67**	**67**	**0**	**44**	**56**	**44**	**44**	**11**

Appendix P Median occurrence rates (%) for the natures and locations of uncertainty within the two (denoted by an asterisk) or three case studies that comprise UnISERA, organised by ERA task and including ERA phase and overall medians, and the highest proportion(s; of at least 50%) for the nature (shaded blue) and location (shaded red) of uncertainty associated with each ERA task (see Section 5.7.1).

ERA Task	Nature of uncertainty (%)			Location of uncertainty (%)						
	Epistemic	Aleatory	Combined	Data	Language	System	Variability	Extrapolation	Model	Decision
1	11	5	79	79	5	63	63	47	47	32
2	26	0	74	84	5	63	68	47	47	26
3	11	5	84	68	11	74	74	42	58	21
4	11	5	84	79	16	53	79	47	63	26
5	37	5	47	53	37	42	32	26	42	11
6	37	5	47	58	32	47	32	26	32	11
7	11	5	58	42	26	37	37	47	53	16
8	21	5	74	74	32	58	47	39	58	28
9	21	5	74	68	21	53	53	53	53	26
10	26	11	47	42	11	42	37	21	21	26
11	32	0	63	63	16	68	42	32	53	16
12	26	5	47	37	11	47	26	11	42	11
13	26	5	58	53	21	53	42	26	32	21
14	21	11	68	53	16	58	47	53	47	16
*15	21	7	71	71	7	36	50	43	36	21
17	16	16	63	42	5	53	42	42	42	26
*18	21	21	50	36	14	36	43	36	29	21
*19	14	14	71	64	21	43	64	57	50	43
22	37	5	58	68	16	63	42	26	37	16
23	11	11	79	47	0	68	79	68	53	32
24	5	5	89	58	32	74	63	63	74	53
*26	14	21	64	64	7	50	79	36	43	14
27	16	11	74	79	16	68	74	37	37	21
28	32	16	63	74	0	58	63	37	37	11
29	21	0	74	68	26	63	32	32	53	53
30	26	0	74	63	26	37	26	32	53	37
31	42	5	47	63	21	26	21	16	21	37
32	42	5	53	63	26	26	21	26	37	37
Problem formulation median	**21**	**5**	**66**	**63**	**16**	**53**	**45**	**37**	**45**	**24**
33	37	5	53	63	21	63	37	16	32	26
34	16	16	63	68	16	63	58	42	53	26
35	16	11	68	68	16	58	47	42	74	37
37	26	11	42	42	16	63	26	37	26	21
*38	29	21	36	50	0	7	57	21	14	0
39	32	26	37	63	16	21	53	26	32	11
40	21	21	53	68	5	32	68	26	32	11
41	21	11	68	84	5	32	68	47	53	16
42	21	11	68	84	5	37	74	47	47	21
43	21	16	63	79	5	32	68	32	16	16
*44	29	29	43	71	0	21	64	29	21	14
45	26	16	53	79	5	32	63	37	32	21
*46	29	14	57	71	0	36	64	29	29	7
*47	29	21	50	71	14	50	64	57	43	14
48	21	11	68	74	5	47	63	47	58	21

49	32	16	47	63	5	16	47	16	21	11
*50	36	7	57	86	0	7	57	7	14	7
52	11	26	53	58	5	11	68	21	26	11
*54	21	21	50	71	0	14	57	43	29	0
55	26	26	42	58	0	21	37	32	26	0
56	16	37	47	47	0	47	58	63	21	11
57	11	21	68	63	0	47	74	63	37	5
58	11	37	53	47	0	32	74	37	37	16
59	16	26	58	58	5	37	63	21	42	11
60	21	16	63	63	0	21	74	42	58	11
*61	42	5	32	42	0	47	21	11	21	11
*64	22	0	67	61	33	72	56	44	67	39
65	21	21	58	47	5	37	47	53	26	21
66	32	16	53	53	5	37	42	47	37	16
67	16	21	58	68	5	26	68	26	26	16
68	11	0	79	47	0	47	47	37	74	21
69	5	0	89	47	0	53	32	53	84	26
Exposure assessment median	**21**	**16**	**55**	**63**	**5**	**36**	**58**	**37**	**32**	**15**
70	47	0	42	74	16	53	32	21	47	5
71	42	5	47	68	5	53	32	32	42	11
*72	21	0	71	71	21	64	50	50	64	14
73	5	16	74	68	5	53	68	74	37	21
74	16	5	74	58	26	21	37	53	47	47
75	21	5	68	74	5	32	53	37	58	11
76	11	26	58	47	5	32	63	47	53	5
77	16	16	68	74	5	26	74	37	53	5
78	11	16	74	74	5	32	74	26	32	16
*79	17	17	67	67	6	22	61	22	17	6
81	11	16	68	63	5	16	68	37	32	5
*82	14	14	50	43	0	14	43	14	14	0
*83	7	14	79	71	7	57	71	57	36	7
*87	0	0	100	79	7	64	71	64	100	36
88	16	5	79	68	0	26	58	53	63	16
89	11	5	79	63	0	26	58	53	63	16
90	11	11	79	63	0	32	63	68	68	16
91	11	0	89	74	0	26	79	63	74	11
Effects assessment median	**12**	**8**	**73**	**68**	**5**	**32**	**62**	**49**	**50**	**11**
92	16	0	84	68	5	53	68	47	58	32
93	16	11	74	68	5	53	68	68	68	32
94	0	11	89	53	11	42	68	79	79	47
95	0	11	89	47	26	42	79	84	79	26
96	0	5	95	47	5	53	63	68	68	47
97	5	0	84	42	5	63	53	63	68	37
100	5	11	84	63	5	37	58	68	42	32
101	5	11	84	58	11	37	58	79	47	26
102	11	11	79	47	42	47	47	53	37	47
103	11	11	79	42	37	58	42	53	42	47
104	5	5	89	58	21	32	58	58	58	42
Risk characterisation median	**5**	**11**	**84**	**53**	**11**	**47**	**58**	**68**	**58**	**37**
Overall median	**19**	**11**	**65**	**62**	**11**	**42**	**55**	**42**	**45**	**21**

Appendix Q ERA tasks in UnISERA organised in order of descending median level of uncertainty, with accompanying ranked occurrence proportions (for median values of at least 50%) for the associated nature and locations of uncertainty (see Section 5.7.2).

ERA Task #	ERA Phase	ERA sub-phase	ERA task group	ERA task	Level[a]	Nature[b]	Location(s)[c]
72	Effects	1. Use available evidence to better constrain…	…	Secondary stressors	7.0 (Ig)	Co	1: Dat; 2: Sys, Mod; 3: Var, Ext;
101	Risk	3. Evaluate risk levels	3.a Assess confidence in the risk levels using…	Experimental evidence	7.0 (Ig)	Co	1: Ext; 2: Dat, Var;
76	Effects	2. Analyse the stressor-response relationship	2.a Determine the test dose for the…	Frequency	6.0 (Sc)	Co	1: Var; 2: Mod;
87	Effects	3. Integrate multiple LOEs using…	…	Quantitative methods	6.0 (Sc)	Co	1: Mod; 2: Dat; 3: Var;
96	Risk	2. Estimate and aggregate risk	2.b Aggregate risk estimates for…	Assessment endpoints	6.0 (Sc)	Co	1: Ext, Mod; 2: Var; 3: Sys;
97	Risk	2. Estimate and aggregate risk	2.b Aggregate risk estimates for…	Stressors	6.0 (Sc)	Co	1: Mod; 2: Sys, Ext; 3: Var;
90	Effects	4. Create stressor-response profile using…	…single point methods showing…	Effects levels	6.0 (Sc)	Co	1: Ext, Mod; 2: Dat, Var;
94	Risk	2. Estimate and aggregate risk	2.a Estimate risk using…	Single-point profiles	6.0 (Sc)	Co	1: Ext, Mod; 2: Var; 3: Dat;
24	Problem	2. Define the conceptual model	2.c Consider the appropriateness of the endpoints	Relative importance of endpoints to each other	6.0 (Sc)	Co	1: Sys, Mod; 2: Var, Ext; 3: Dat;
89	Effects	4. Create stressor-response profile using…	…single point methods showing…	Extreme toxicity	6.0 (Sc)	Co	1: Dat, Mod; 2: Var; 3: Ext;
103	Risk	3. Evaluate risk levels	3.b Assess the significance of the risk levels using…	Stakeholder levels	6.0 (Sc)	Co	1: Sys; 2: Ext;
68	Exposure	5. Create the exposure profile(s) using…	…	Stressor-based models	5.0 (Sc)	Co	1: Mod;
93	Risk	1. Select relevant profiles…	…distribution methods showing	For effects	5.0 (Sc)	Co	1: Dat, Var, Ext, Mod; 2: Sys;
2	Problem	1. Preliminary hazard identification	1.a Use available evidence to better constrain…	Potential receptors	5.0 (Sc)	Co	1: Dat; 2: Var; 3: Sys;
60	Exposure	3. Evaluate stressor-receptor contact	3.a Evaluate co-occurrence for…	Intensity of overlap	5.0 (Sc)	Co	1: Var; 2: Dat; 3: Mod;

102	Risk	3. Evaluate risk levels	3.b Assess the significance of the risk levels using…	Regulatory levels	5.0 (Sc)	Co	1: Ext;
64	Exposure	4. Integrate multiple LOEs using…	…	Quantitative methods	5.0 (Sc)	Co	1: Sys; 2: Mod; 3: Dat;
100	Risk	3. Evaluate risk levels	3.a Assess confidence in the risk levels using…	Empirical evidence	5.0 (Sc)	Co	1: Ext; 2: Dat; 3: Var;
104	Risk	3. Evaluate risk levels	3.b Assess the significance of the risk levels using…	Experimental levels	5.0 (Sc)	Co	1: Dat, Var, Ext, Mod;
69	Exposure	5. Create the exposure profile(s) using…	…	Receptor-based models	5.0 (Sc)	Co	1: Mod; 2: Sys, Ext;
4	Problem	1. Preliminary hazard identification	1.a Use available evidence to better constrain…	Potential effects	5.0 (Sc)	Co	1: Dat, Var; 2: Mod; 3: Sys;
42	Exposure	2. Stressor, exposure media, and receptor information	2.a.2 Collect information about the stressor's distribution	Temporal	5.0 (Sc)	Co	1: Dat; 2: Var;
35	Exposure	1. Use available evidence to better constrain…	…	Model selection	5.0 (Sc)	Co	1: Mod; 2: Dat; 3: Sys;
57	Exposure	2. Stressor, exposure media, and receptor information	2.c Collect information about the receptor	Temporal distribution	5.0 (Sc)	Co	1: Var; 2: Dat, Ext;
19	Problem	2. Define the conceptual model	2.b Choose assessment and measurement endpoints	Population: abundance	5.0 (Sc)	Co	1: Dat, Var; 2: Ext;
48	Exposure	2. Stressor, exposure media, and receptor information	2.b Collect information about properties affecting fate and transport	Physical	5.0 (Sc)	Co	1: Dat; 2: Var; 3: Mod;
66	Exposure	5. Create the exposure profile(s) using…	…	Worst-case estimates	5.0 (Sc)	Co	1: Dat;
91	Effects	4. Create stressor-response profile using…	…distribution methods showing	Effects levels	5.0 (Sc)	Co	1: Var; 2: Dat, Mod; 3: Ext;
75	Effects	2. Analyse the stressor-response relationship	2.a Determine the test dose for the…	Duration	5.0 (Sc)	Co	1: Dat; 2: Mod; 3: Var;
22	Problem	2. Define the conceptual model	2.c Consider the appropriateness of the endpoints	Relevance of measures to their endpoints	5.0 (Sc)	Co	1: Dat; 2: Sys;
95	Risk	2. Estimate and aggregate risk	2.a Estimate risk using…	Cumulative distributions	5.0 (Sc)	Co	1: Ext; 2: Var, Mod;
71	Effects	1. Use available evidence to better constrain…	…	Direct/indirect effects	4.5 (Sc)	-	1: Dat; 2: Sys;
83	Effects	2. Analyse the stressor-response relationship	2.b Assess effect endpoints	Population: abundance	4.5 (Sc)	Co	1: Dat, Var; 2: Sys, Ext;
61	Exposure	3. Evaluate stressor-receptor contact	3.b. Evaluate…	Nature of contact	4.0 (Sc)	-	-

67	Exposure	5. Create the exposure profile(s) using…	…	Direct monitoring values	4.0 (Sc)	Co	1: Dat, Var;
56	Exposure	2. Stressor, exposure media, and receptor information	2.c Collect information about the receptor	Spatial distribution	4.0 (Sc)	-	1: Ext; 2: Var;
23	Problem	2. Define the conceptual model	2.c Consider the appropriateness of the endpoints	Significance of endpoints to receptor	4.0 (Sc)	Co	1: Var; 2: Sys, Ext; 3: Mod;
3	Problem	1. Preliminary hazard identification	1.a Use available evidence to better constrain…	Potential exposure	4.0 (Sc)	Co	1: Sys, Var; 2: Dat; 3: Mod;
92	Risk	1. Select relevant profiles…	…distribution methods showing	For exposure	4.0 (Sc)	Co	1: Dat, Var; 2: Mod; 3: Sys;
73	Effects	1. Use available evidence to better constrain…	…	Toxicity levels	4.0 (Sc)	Co	1: Ext; 2: Dat, Var; 3: Sys;
29	Problem	3. Form the analysis/work plan	3.b Identify data considerations	Gaps in data	4.0 (Sc)	Co	1: Dat; 2: Sys; 3: Mod, Dec;
65	Exposure	5. Create the exposure profile(s) using…	…	Conservative estimates	4.0 (Sc)	Co	1: Ext;
15	Problem	2. Define the conceptual model	2.b Choose assessment and measurement endpoints	Organism: behaviour	4.0 (Sc)	Co	1: Dat;
37	Exposure	1. Use available evidence to better constrain…	…	Prioritisation of data	4.0 (Sc)	-	1: Sys;
79	Effects	2. Analyse the stressor-response relationship	2.b Assess effect endpoints	Organism: behaviour	4.0 (Sc)	Co	1: Dat; 2: Var;
59	Exposure	3. Evaluate stressor-receptor contact	3.a Evaluate co-occurrence for…	Temporal overlap	4.0 (Sc)	Co	1: Var; 2: Dat;
58	Exposure	3. Evaluate stressor-receptor contact	3.a Evaluate co-occurrence for…	Spatial overlap	4.0 (Sc)	Co	1: Var;
14	Problem	2. Define the conceptual model	2.b Choose assessment and measurement endpoints	Organism: development	4.0 (Sc)	Co	1: Sys; 2: Dat, Ext;
34	Exposure	1. Use available evidence to better constrain…	…	Exposure levels	4.0 (Sc)	Co	1: Dat; 2: Sys; 3: Var;
33	Exposure	1. Use available evidence to better constrain…	…	Nature of exposure	4.0 (Sc)	Co	1: Dat, Sys;
78	Effects	2. Analyse the stressor-response relationship	2.b Assess effect endpoints	Organism: development	4.0 (Sc)	Co	1: Dat, Var;
46	Exposure	2. Stressor, exposure media, and receptor information	2.b Collect information about properties affecting fate and transport	Biological	4.0 (Sc)	Co	1: Dat, 2: Var;
27	Problem	3. Form the analysis/work plan	3.a Identify the factors controlling fate and transport of the stressor	Physical factors	4.0 (Sc)	Co	1: Dat; 2: Var; 3: Sys;

88	Effects	4. Create stressor-response profile using…	…single point methods showing…	Conservative toxicity	4.0 (Sc)	Co	1: Dat; 2: Mod; 3: Var;
74	Effects	1. Use available evidence to better constrain…	…	Prioritisation of data	4.0 (Sc)	Co	1: Dat; 2: Ext;
70	Effects	1. Use available evidence to better constrain…	…	Nature of effects	3.5 (Sc)	-	1: Dat; 2: Sys;
47	Exposure	2. Stressor, exposure media, and receptor information	2.b Collect information about properties affecting fate and transport	Chemical	3.5 (Sc)	Co	1: Dat; 2: Var; 3: Ext;
26	Problem	3. Form the analysis/work plan	3.a Identify the factors controlling fate and transport of the stressor	Chemical factors	3.5 (Sc)	Co	1: Var; 2: Dat;
1	Problem	1. Preliminary hazard identification	1.a Use available evidence to better constrain…	Potential stressors	3.5 (Sc)	Co	1: Dat; 2: Sys, Var;
55	Exposure	2. Stressor, exposure media, and receptor information	2.c Collect information about the receptor	Receptor characteristics	3.1 (St)	-	1: Dat;
43	Exposure	2. Stressor, exposure media, and receptor information	2.a.3 Collect information info about the stressor's release	Intensity	3.0 (St)	Co	1: Dat; 2: Var;
32	Problem	3. Form the analysis/work plan	3.b Identify data considerations	Analysis techniques	3.0 (St)	Co	1: Dat;
18	Problem	2. Define the conceptual model	2.b Choose assessment and measurement endpoints	Organism: fecundity	3.0 (St)	Co	-
30	Problem	3. Form the analysis/work plan	3.b Identify data considerations	Types of data required	3.0 (St)	Co	1: Dat; 2: Mod;
54	Exposure	2. Stressor, exposure media, and receptor information	2.c Collect information about the receptor	Physical composition	3.0 (St)	Co	1: Dat; 2: Var;
50	Exposure	2. Stressor, exposure media, and receptor information	2.b Collect information about properties affecting fate and transport	Environmental media: biota	3.0 (St)	Co	1: Dat; 2: Var;
9	Problem	1. Preliminary hazard identification	1.b Framing the hazard	Frame the 'when'	3.0 (St)	Co	1: Dat; 2: Sys, Var, Ext, Mod;
82	Effects	2. Analyse the stressor-response relationship	2.b Assess effect endpoints	Organism: fecundity	3.0 (St)	Co	-
45	Exposure	2. Stressor, exposure media, and receptor information	2.a.3 Collect information info about the stressor's release	Quantity	3.0 (St)	Co	1: Dat; 2: Var;
49	Exposure	2. Stressor, exposure media, and receptor information	2.b Collect information about properties affecting fate and transport	Environmental media: terrestrial	3.0 (St)	-	1: Dat;
41	Exposure	2. Stressor, exposure media, and receptor information	2.a.2 Collect information about the stressor's distribution	Spatial	3.0 (St)	Co	1: Dat; 2: Var; 3: Mod;
17	Problem	2. Define the conceptual model	2.b Choose assessment and measurement endpoints	Organism: survival	3.0 (St)	Co	1: Sys;
28	Problem	3. Form the analysis/work plan	3.a Identify the factors controlling fate and transport of the stressor	Environmental media factors	3.0 (St)	Co	1: Dat; 2:

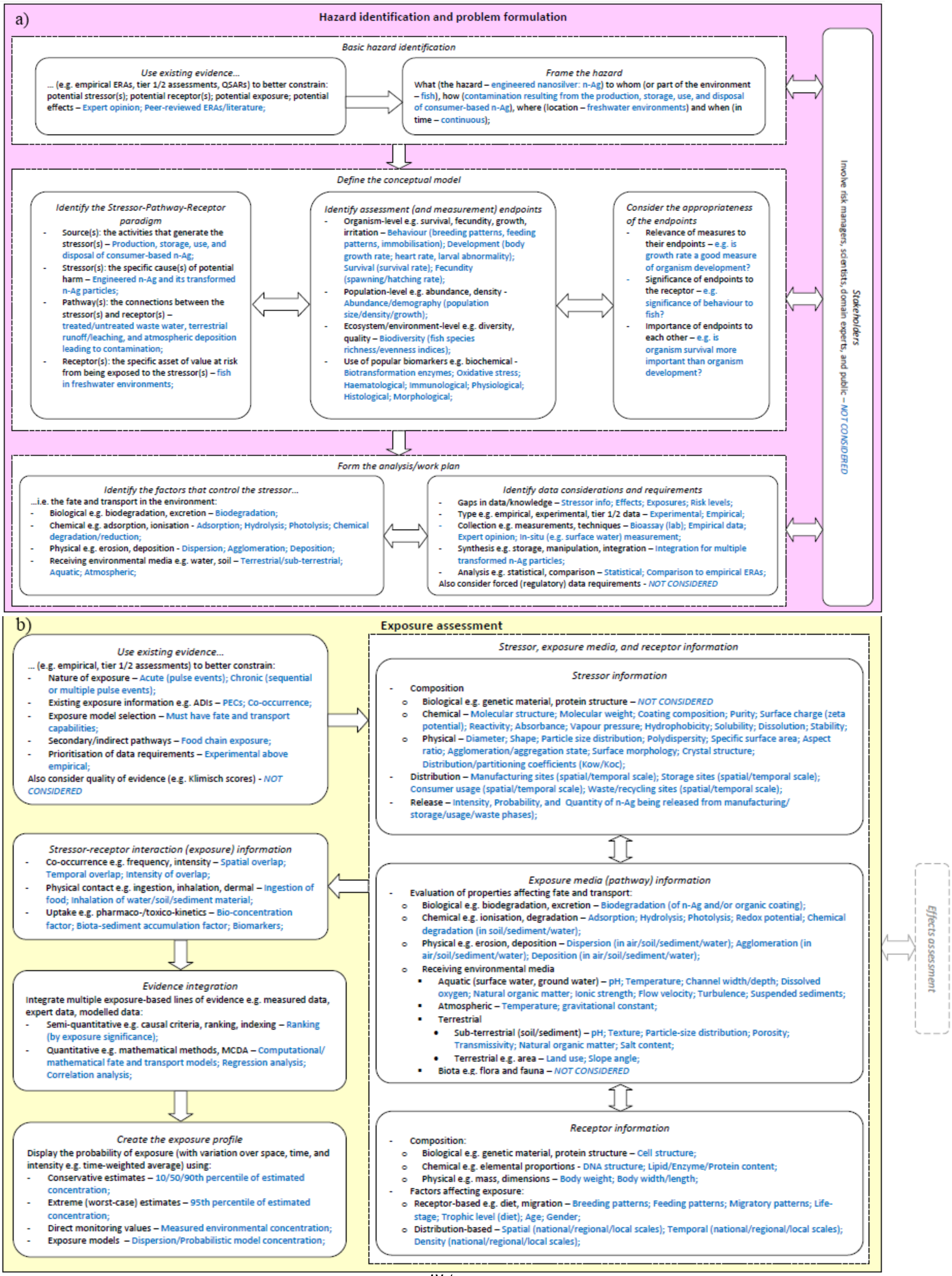
							Var; 3: Sys;
38	Exposure	2. Stressor, exposure media, and receptor information	2.a.1 Collect information about the stressor's composition	Biological information	3.0 (St)	-	1: Var;
8	Problem	1. Preliminary hazard identification	1.b Framing the hazard	Frame the 'where'	3.0 (St)	Co	1: Dat; 2: Sys, Mod;
52	Exposure	2. Stressor, exposure media, and receptor information	2.b Collect information about properties affecting fate and transport	Environmental media: Atmospheric	3.0 (St)	Co	1: Var; 2: Dat;
44	Exposure	2. Stressor, exposure media, and receptor information	2.a.3 Collect information info about the stressor's release	Probability	3.0 (St)	-	1: Dat; 2: Var;
40	Exposure	2. Stressor, exposure media, and receptor information	2.a.1 Collect information about the stressor's composition	Physical information	3.0 (St)	Co	1: Dat, Var;
81	Effects	2. Analyse the stressor-response relationship	2.b Assess effect endpoints	Organism: survival	3.0 (St)	Co	1: Var; 2: Dat;
77	Effects	2. Analyse the stressor-response relationship	2.a Determine the test dose for the…	Intensity	3.0 (St)	Co	1: Dat, Var; 2: Mod;
7	Problem	1. Preliminary hazard identification	1.b Framing the hazard	Frame the 'how'	2.5 (St)	Co	1: Mod;
6	Problem	1. Preliminary hazard identification	1.b Framing the hazard	Frame the 'whom'	2.0 (St)	-	1: Dat;
31	Problem	3. Form the analysis/work plan	3.b Identify data considerations	Collection techniques	2.0 (St)	-	1: Dat;
13	Problem	2. Define the conceptual model	2.a Identify the S-P-R paradigm, including…	The receptor(s)	2.0 (St)	Co	1: Dat, Sys;
39	Exposure	2. Stressor, exposure media, and receptor information	2.a.1 Collect information about the stressor's composition	Chemical information	2.0 (St)	-	1: Dat; 2: Var;
10	Problem	2. Define the conceptual model	2.a Identify the S-P-R paradigm, including…	The source(s)	2.0 (St)	-	-
11	Problem	2. Define the conceptual model	2.a Identify the S-P-R paradigm, including…	The stressor(s)	2.0 (St)	Co	1: Sys; 2: Dat; 3: Mod;
12	Problem	2. Define the conceptual model	2.a Identify the S-P-R paradigm, including…	The pathway(s)	2.0 (St)	-	-
5	Problem	1. Preliminary hazard identification	1.b Framing the hazard	Frame the 'what'	1.5 (St)	-	1: Dat;

[a] Ig=Recognised ignorance; Sc=Scenario uncertainty; St=Statistical uncertainty. Statistical significance (*P*) is used to rank like values.

[b] Co=Combined.

[c] Dat=Data; Lan=Language; Sys=System; Var=Variability; Ext=Extrapolation; Mod=Model; Dec=Decision. Median occurrence proportions are used to rank like values.

Appendix R The consumer-based engineered nano-Ag risk to freshwater fish ERA template, version 1, created by populating the generic ERA template, version 3, with information from relevant articles, describing the aspects within the stages of: **a)** hazard identification and problem formulation; **b)** exposure assessment; **c)** effects assessment; and **d)** risk characterisation (see Section 6.3.2).



c) **Effects assessment (A.K.A. Stressor/Dose/Concentration-response)**

Use existing evidence…

… (e.g. empirical, tier 1/2 assessments) to better constrain:

- Potential direct/indirect/secondary effects – (see endpoints);
- Nature of effects e.g. internal/external – Internal; ADME data;
- Potential secondary/indirect stressors – Transformed n-Ag;
- Existing toxicity levels – $LC_{10/50}$; EC_{50}; NOAEL; PNEC; Species sensitivity distributions (SSDs); Water quality criteria (WQC);
- Prioritisation of data requirements - Experimental above empirical;

Also consider quality of evidence e.g. Klimisch scores - *NOT CONSIDERED*

Evidence integration

Integrate multiple effects-based lines of evidence e.g. measured data and modelled data:

- Semi-quantitative e.g. causal criteria, ranking, indexing – Ranking (by effects significance);
- Quantitative e.g. mathematical methods/models, MCDA – Mathematical models; Regression analysis; Correlation analysis; Spatially explicit models;

Create the stressor-response profile(s)

For each assessment endpoint display the effects and their magnitudes (with variation over space, time, and intensity) using either:

- Single-point profiles:
  - Conservative toxicity (i.e. acute) levels - LC_{50}; EC_{50};
  - Extreme toxicity (i.e. acute) levels - LC_{10};
  - Effects (i.e. chronic) levels e.g. NOAEL, LOAEL – NOAEL; PNEC (NOAEL / assessment factor);
- Distribution profiles
  - Species sensitivity (SSD) - lognormal of toxicity or effects profiles;
  - Stressor-response (i.e. acute and/or chronic) – distribution (of intensity of stressor Vs. magnitude of response);

Analyse the stressor-response relationship

The test dose(s)

Quantify the internal/external test dose (of the stressor) received (by each receptor) via the pathway(s):

- Concentration – of test dose;
- Duration – of exposure to test dose;
- Frequency (number of applications) – of test dose;
- Quantity – of n-Ag in water sample;

Re-evaluate for multiple stressors – Transformed n-Ag particles

Assess the effect endpoints…

…using the measurement endpoints:

Assessment endpoint	Measure
Organism behaviour	Breeding patterns; Feeding patterns; Immobilisation
Organism development	Body growth rate; Heart rate; Larval abnormality
Organism survival	Survival rate
Organism fecundity	Spawning rate; Hatching rate
Organism biomarkers	Biotransformation enzymes; Oxidative stress; Haematological; Immunological; Physiological; Histological; Morphological
Population abundance/demography	Population size/density/growth
Ecosystem biodiversity	Fish species richness/evenness indices

… and outlined data collection techniques – Bioassay (lab); Empirical data; Expert opinion; In-situ measurement;

Also consider:

- Criteria for establishing causality between stressor(s) and effect(s) e.g. strength, consistency, predictability – Strength of association; Consistency of association;
- Extrapolations required e.g. time, space – Within/between species; Time; Space; Quantity; Lab to field;
- Secondary/indirect effects – Changes in water quality (WQC);

Exposure assessment

d) **Risk characterisation**

Consider relevant exposure and effect profiles

Relevant profiles are those that:

- For exposure, display any exposure(s) of receptor to stressor;
- For effect, display any effect(s) of stressor upon receptor;

Risk estimation and aggregation

Estimate the risk

For each assessment endpoint evaluated, *combine* the effect profile with the exposure profile (describing variation over space, time, and/or intensity) using either:

- Single point comparison:
  - Risk quotient : PEC / LC_{50}; PEC / EC_{50}; PEC / NOAEL; PEC / PNEC;
  - Quality criterion: HC_5 value of a SSD;
- Cumulative/Frequency distribution:
  - Risk for individual/multiple n-Ag transformation (s):
    - Expected ecological risk: Exposure curve Vs. stressor-response curve;
    - Potentially affected fraction (PAF) = SSD + PEC;

Aggregate risk estimates

Where necessary and depending on the focus (i.e. single/multiple stressor(s)/pathway(s)/receptor(s)), consider aggregating relevant risk estimates for:

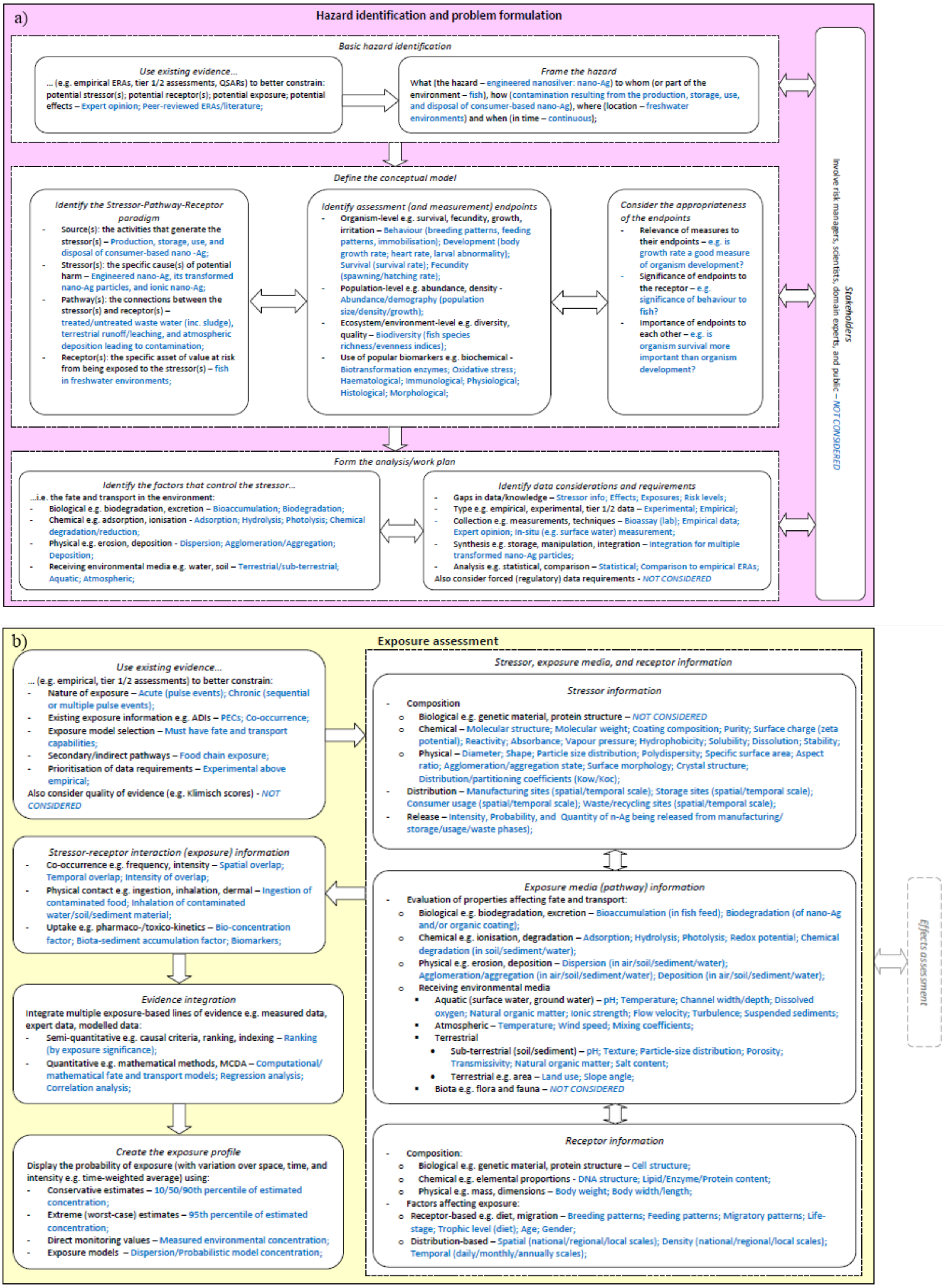
- Assessment endpoints – Behaviour; Development; Survival; Fecundity; Abundance/demography; Biodiversity;
- Stressors – Multiple transformed n-Ag particles;
- Pathways – Treated/untreated waste water; Terrestrial runoff/leaching; Atmospheric deposition;
- Receptors – Multiple fish species;

…using previously outlined evidence integration techniques.

Risk evaluation

- Confidence in the risk estimate(s), based on:
  - Uncertainty analysis – *NOT CONSIDERED (in the context of this template validation exercise, at least)*
    - Semi-quantitative e.g. Assessment factors
    - Quantitative e.g. Monte-Carlo
  - Supporting empirical (i.e. peer-reviewed) evidence – e.g. Lab-based; Field-based; Model-based; Expert opinion;
  - Supporting experimental (i.e. non peer-reviewed) evidence – e.g. Lab-based; Field-based; Model-based; Expert opinion;
- Significance of the risk level, based on:
  - Regulatory-enforced acceptability levels – i.e. Expert-derived and enforced by appropriate regulator(s);
  - Stakeholder-derived acceptability levels – i.e. Those involved in current risk assessment;
  - Experimentally-derived acceptability levels – i.e. Other risk assessments and relevant academic/grey research;
  - Receptor recovery potential – Recovery from acute low-level exposure;

Appendix S The consumer-based engineered nano-Ag risk to freshwater fish ERA template, version 2, created through the expert validation of version 1, describing the important aspects within the phases of: **a)** hazard identification and problem formulation; **b)** exposure assessment; **c)** effects assessment; and **d)** risk characterisation (see Section 6.3.2).



c)

Effects assessment (A.K.A. Stressor/Dose/Concentration-response)

Use existing evidence…

… (e.g. empirical, tier 1/2 assessments) to better constrain:

- Potential direct/indirect/secondary effects – (see endpoints);
- Nature of effects e.g. internal/external – Internal; ADME data;
- Potential secondary/indirect stressors – Transformed nano-Ag;
- Existing toxicity levels – $LC_{10/50}$; EC_{50}; NOAEL; PNEC; Benchmark Dose (BMD); Species sensitivity distributions (SSDs); Water quality criteria (WQC);
- Prioritisation of data requirements - Experimental above empirical;

Also consider quality of evidence e.g. Klimisch scores - *NOT CONSIDERED*

Analyse the stressor-response relationship

The test dose(s)

Quantify the internal/external test dose (of the stressor) received (by each receptor) via the pathway(s):

- Concentration – of test dose;
- Duration – of exposure to test dose;
- Frequency (number of applications) – of test dose;
- Quantity – of nano-Ag in water sample;

Re-evaluate for multiple stressors – Transformed nano-Ag particles

Assess the effect endpoints…

…using the measurement endpoints:

Assessment endpoint	Measure
Organism behaviour	Breeding patterns; Feeding patterns; Immobilisation
Organism development	Body growth rate; Heart rate; Larval abnormality
Organism survival	Survival rate
Organism fecundity	Spawning rate; Hatching rate
Organism biomarkers	Biotransformation enzymes; Oxidative stress; Haematological; Immunological; Physiological; Histological; Morphological
Population abundance/demography	Population size/density/growth
Ecosystem biodiversity	Fish species richness/evenness indices

… and outlined data collection techniques – Bioassay (lab); Empirical data; Expert opinion; In-situ measurement;

Also consider:

- Criteria for establishing causality between stressor(s) and effect(s) e.g. strength, consistency, predictability – Strength of association; Consistency of association;
- Extrapolations required e.g. time, space – Within/between species; Time; Space; Quantity; Lab to field;
- Secondary/indirect effects – Changes in water quality (WQC);

Evidence integration

Integrate multiple effects-based lines of evidence e.g. measured data and modelled data:

- Semi-quantitative e.g. causal criteria, ranking, indexing – Ranking (by effects significance);
- Quantitative e.g. mathematical methods/models, MCDA – Mathematical models; Regression analysis; Correlation analysis; Spatially explicit models;

Create the stressor-response profile(s)

For each assessment endpoint display the effects and their magnitudes (with variation over space, time, and intensity) using either:

- Single-point profiles:
    - Conservative toxicity (i.e. acute) levels - LC_{50}; EC_{50};
    - Extreme toxicity (i.e. acute) levels - LC_{10};
    - Effects (i.e. chronic) levels e.g. NOAEL, LOAEL – NOAEL; PNEC (NOAEL / assessment factor); Benchmark Dose (BMD);
- Distribution profiles
    - Species sensitivity (SSD) - lognormal of toxicity or effects profiles;
    - Stressor-response (i.e. acute and/or chronic) – distribution (of intensity of stressor Vs. magnitude of response);

Exposure assessment

d)

Risk characterisation

Consider relevant exposure and effect profiles

Relevant profiles are those that:

- For exposure, display any exposure(s) of receptor to stressor;
- For effect, display any effect(s) of stressor upon receptor;

Risk estimation and aggregation

Estimate the risk

For each assessment endpoint evaluated, *combine* the effect profile with the exposure profile (describing variation over space, time, and/or intensity) using either:

- Single point comparison:
    - Risk quotient : PEC / LC_{50}; PEC / EC_{50}; PEC / NOAEL; PEC / PNEC;
    - Quality criterion: HC_5 value of a SSD;
- Cumulative/Frequency distribution:
    - Risk for individual/multiple n-Ag transformation (s):
        - Expected ecological risk: Exposure curve Vs. stressor-response curve;
        - Potentially affected fraction (PAF) = SSD + PEC;

Aggregate risk estimates

Where necessary and depending on the focus (i.e. single/multiple stressor(s)/pathway(s)/receptor(s)), consider aggregating relevant risk estimates for:

- Assessment endpoints – Behaviour; Development; Survival; Fecundity; Abundance/demography; Biodiversity;
- Stressors – Multiple transformed nano-Ag particles;
- Pathways – Treated/untreated waste water; Terrestrial runoff/leaching; Atmospheric deposition;
- Receptors – Multiple fish species;

…using previously outlined evidence integration techniques.

Risk evaluation

- Confidence in the risk estimate(s), based on:
    - Uncertainty analysis – *NOT CONSIDERED (in the context of this template validation exercise, at least)*
        - Semi-quantitative e.g. Assessment factors
        - Quantitative e.g. Monte-Carlo
    - Supporting empirical (i.e. peer-reviewed) evidence – e.g. Lab-based; Field-based; Model-based; Expert opinion;
    - Supporting experimental (i.e. non peer-reviewed) evidence – e.g. Lab-based; Field-based; Model-based; Expert opinion;
- Significance of the risk level, based on:
    - Regulatory-enforced acceptability levels – i.e. Expert-derived and enforced by appropriate regulator(s);
    - Stakeholder-derived acceptability levels – i.e. Those involved in current risk assessment;
    - Experimentally-derived acceptability levels – i.e. Other risk assessments and relevant academic/grey research;
    - Receptor recovery potential – Recovery from acute low-level exposure;

Appendix T Median occurrence rates (%) for the natures and locations of uncertainty within the Validation Case Study, organised by ERA task and including ERA phase and overall medians, and the highest proportion(s; at least 50%) for the nature (shaded blue) and location (shaded red) of uncertainty associated with each ERA task (see Section 6.3.3).

ERA Task	**Nature of uncertainty (%)**			**Location of uncertainty (%)**						
	Epistemic	**Aleatory**	**Combined**	**Data**	**Language**	**System**	**Variability**	**Extrapolation**	**Model**	**Decision**
1	17	17	50	17	17	33	33	50	33	0
2	50	0	50	67	0	67	33	33	33	0
3	17	0	83	83	17	83	67	33	83	17
4	33	17	50	17	17	83	17	17	0	0
5	33	17	33	67	50	33	17	17	17	17
6	17	33	33	33	33	17	50	0	0	17
7	17	0	83	100	0	83	50	33	83	0
8	33	0	33	83	0	50	33	17	17	0
9	0	33	67	67	0	67	100	17	50	0
10	17	17	33	67	17	17	67	33	33	0
11	50	0	50	67	33	100	0	50	17	17
12	33	0	67	83	17	83	33	17	0	0
13	17	17	67	67	0	17	50	33	0	17
14	0	0	100	83	17	50	50	33	0	17
15	17	17	50	83	17	33	50	33	0	0
17	0	0	100	83	17	83	50	33	17	0
18	0	0	100	83	17	83	83	33	0	17
19	0	0	100	100	17	83	100	17	50	17
21	0	0	100	83	17	83	83	17	17	17
22	0	17	83	33	0	83	67	67	33	0
23	33	0	67	67	0	50	67	33	17	0
24	17	0	83	67	33	67	67	33	33	17
25	67	17	17	67	17	50	17	33	17	0
26	33	17	17	33	33	33	17	17	17	0
27	0	0	100	83	50	83	83	0	50	0
28	67	0	33	67	17	100	33	17	0	0
29	0	0	83	67	50	50	67	83	33	0
30	83	0	17	67	17	50	17	0	17	0
31	33	0	67	50	33	83	50	17	33	0
32	50	0	50	33	17	83	50	0	17	17
Problem formulation median	**17**	**0**	**67**	**67**	**17**	**67**	**50**	**33**	**17**	**0**
33	17	0	83	100	0	83	67	33	33	0
34	17	0	83	100	0	17	33	50	17	0
35	17	0	83	83	33	67	50	50	50	0
36	33	0	67	100	0	50	67	67	33	0
37	33	17	33	67	33	50	17	17	17	17
39	67	17	17	33	17	50	33	0	17	0
40	67	17	17	50	17	50	33	0	17	0
41	33	17	33	83	33	0	67	17	17	0
42	17	0	83	100	33	50	83	17	17	0
43	17	0	83	100	0	33	83	17	50	0
44	17	0	83	100	17	83	83	33	67	0
45	50	0	50	100	17	17	50	33	17	0
46	17	0	83	100	33	83	67	33	50	0

47	50	0	50	83	17	50	50	33	33	0
48	50	0	50	83	33	50	50	17	0	0
49	33	17	33	83	0	0	67	0	0	0
51	33	17	33	83	0	0	67	0	0	0
52	17	17	67	83	0	0	67	17	0	0
53	17	17	67	83	0	33	83	17	0	0
54	33	17	33	83	0	0	50	17	0	0
55	17	17	50	83	0	17	67	17	0	0
56	17	17	67	83	0	0	67	17	0	0
57	17	17	67	83	0	0	83	17	0	0
58	17	0	83	83	0	83	83	33	67	0
59	17	0	83	67	0	50	67	33	83	0
60	17	0	83	100	0	67	67	33	83	0
61	17	17	67	100	0	50	67	17	67	0
62	17	0	83	100	0	67	83	17	67	0
63	33	0	67	100	33	83	33	33	67	17
64	17	0	83	100	17	50	83	17	83	33
65	33	0	67	100	17	50	67	33	17	0
66	67	0	33	83	17	33	33	17	17	0
67	50	0	50	100	0	17	50	50	17	0
68	50	0	50	100	0	83	50	0	50	17
69	17	0	83	83	17	67	83	33	50	17
Exposure assessment median	**17**	**0**	**67**	**83**	**0**	**50**	**67**	**17**	**17**	**0**
70	67	0	33	83	50	67	33	0	0	0
71	67	0	33	83	33	50	33	17	0	0
72	17	0	83	100	0	17	50	50	17	17
73	17	0	83	100	17	33	50	17	33	0
74	33	0	67	83	0	67	67	33	17	33
75	67	0	33	100	0	83	33	17	17	0
76	33	0	67	100	0	83	67	17	17	0
77	67	0	33	100	0	83	33	17	17	0
78	0	17	83	67	17	33	67	67	17	17
79	0	17	83	67	17	33	67	67	17	17
81	33	17	33	67	0	17	17	67	17	17
82	0	17	83	67	0	33	67	83	17	17
83	0	17	83	67	0	67	33	83	17	17
85	0	17	83	67	0	67	67	100	17	33
86	17	0	83	67	0	67	67	33	67	17
87	17	0	83	83	0	67	67	33	67	17
88	17	0	83	83	0	83	83	17	17	0
89	17	0	83	33	17	33	67	83	17	0
90	33	0	67	67	33	67	67	50	17	0
91	33	0	67	83	17	67	67	17	50	0
Effects assessment median	**17**	**0**	**83**	**83**	**0**	**67**	**67**	**33**	**17**	**8**
92	17	17	67	33	0	50	50	33	17	50
93	17	17	67	33	0	50	50	33	17	50
94	0	0	100	67	17	33	67	83	33	17
95	0	0	100	67	0	33	67	83	17	0
96	17	0	83	33	0	67	33	50	0	50
97	17	0	83	33	17	83	33	50	17	50
98	17	0	83	67	17	67	33	50	17	50
99	17	0	83	67	17	67	33	50	17	50
100	50	0	50	83	33	17	33	33	0	17
101	0	0	100	83	33	17	83	67	33	33
102	33	0	67	17	50	83	0	0	17	67
103	33	0	67	17	50	83	0	0	0	67
104	33	0	67	33	50	17	17	33	0	50
105	33	0	67	100	0	17	67	33	0	0

Risk characterisation median	**17**	**0**	**83**	**33**	**17**	**50**	**33**	**50**	**17**	**50**
Overall median	**17**	**0**	**67**	**83**	**17**	**50**	**50**	**33**	**17**	**0**